A design and evaluation of a miniaturised position-sensitive energetic particle detector for small satellites

Arrow Lee

Mullard Space Science Laboratory

UCL

A thesis submitted for the degree of

Doctor of Philosophy

I, Arrow Lee, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis is concerned with the design and evaluation of an *in-situ* energetic particle detector for space weather measurements consisting of a position-sensitive solid state detector and a coded aperture mask. The combination of the two presents a novel technique that provides far higher resolution angular information of supra-thermal particles in low-flux environments in two dimensions than previously possible, over a wide field of view.

After consideration of the historical and current status of *in-situ* plasma instruments and coded aperture techniques, a proposed design was modelled with the Geant4 toolkit.

A proof-of-concept instrument has been designed and assembled and then tested in the laboratory with three electron sources and the results have been analysed. While the available hardware limited the amount of characterisation possible, the instrument made impressive strides in understanding the setup itself and demonstrated its potential within the tested energy ranges.

Further to this, a simulation of such an instrument or its descendants in a space plasma environment was completed and an outlook on the use of the concept in space was discussed. In some configurations this will allow simultaneous energy and high-resolution angular measurements of energetic particle fluxes on small satellites. If such an instrument were flown on a swarm mission, the possibilities for multi-point directional measurements of energetic particles would be hugely increased, especially at low altitudes.

Contents

List of Figures

List of Tables

List of Abbreviations

ACE Advanced Composition Explorer 45

ADC analogue-to-digital converter 118, 129

ADU Analogue to Digital Units 18, 236, 249, 253, 256, 257

AGW Acoustic Gravitational Wave 50

AIM Asteroid Impact Mission 72

AMPS Advanced Miniature Plasma Spectrometer 89

AMPTE Active Magnetospheric Particle Tracer Explorers 82

APS active pixel sensor 86

ASIC application-specific integrated circuit 86, 88

ASM All-Sky Monitor 152

CAD Computer-Aided Design 15, 156, 159

CATS Conceptual And Tiny Spectrometer 87, 119, 276

CCD charge coupled device 15, 17–19, 34, 89, 108, 111–113, 124, 134, 139, 143, 203, 207, 209–213, 217–225, 230–232, 234–237, 239, 240, 244–247, 249, 251, 253, 255, 256, 258, 262, 264, 267, 268, 271, 272, 280–282, 292

CCE charge collection efficiency 174, 251

CDHS Command and Data Handling System 138

CEM Channel electron multiplier 8, 79, 89, 229

CERN European Organization for Nuclear Research 125, 158

CINEMA Cubesat for Ion, Neutral, Electron, Magnetic fields 14, 89–91

CIR co-rotating interaction region 43, 44

CME coronal mass ejection 40, 44–46

CMOS complementary metal-oxide-semiconductor 111, 113, 124, 125, 129, 134, 139, 209, 289

COTS Commercial off-the-shelf 84

CRRES Combined Release and Radiation Effects Satellite 52

CSSWE Colorado Student Space Weather Experiment 86, 91

CSV comma-separated variables 163, 174, 249

EEJ Equatorial Electrojet 57

EISCAT European Incoherent Scatter 67

ELFIN Electron Loss and Field Investigator 140

EM-1 Exploration Mission 1 72

EPD Energetic Particles Detector 65

ERNE Energetic and Relativistic Nuclei and Electron instrument 80

ESA European Space Agency 64, 65, 72, 84, 109, 154

ESTAR STopping power And Range tables for electrons 171, 180, 208

FCFOV fully-coded field of view 155, 190, 199, 201, 204

FFT fast Fourier transform 151

FOV Field of View 140, 145, 148, 154, 168, 176, 188, 190, 283, 290, 293, 295, 296

FWHM full-width half max 154

GCR Galactic Cosmic Ray 14, 59, 60

GDML Geometry Description Markup Language 159

GM Geiger-Müller 76

GNSS Global Navigation Satellite Systems 32, 55

GOCE Gravity field and steady-state Ocean Circulation Explorer 69

GOES Geostationary Operational Environmental Satellite 209

GPS Global Positioning System 67, 68

GRAS Geant4 Radiation Analysis for Space 115

GRB Gamma Ray Burst 154

GRIS Gamma-Ray Imaging Spectrometer 152

HEP High Energy Particle instrument 64

HMRM Highly Miniaturised Radiation Monitor 14, 86–88, 91

HURA Hexagonal URA 148

HV high voltage 222, 225, 226, 231, 243, 276

IBIS Imager on Board of the INTEGRAL Satellite 15, 154–156

IC Integrated Circuit 281

IDL Interactive Data Language 212, 231

InSight Interior Exploration using Seismic Investigations, Geodesy and Heat Transport 72

INSPIRE Interplanetary NanoSpacecraft Pathfinder In Relevant Environment 72

INTEGRAL INTErnational Gamma-Ray Astrophysics Laboratory 29, 154, 155, 212

IRM Ion Release Module 82

ISA Industry Standard Architecture 222, 276

ISIS Innovative Solutions In Space 73

ISS International Space Station 47, 69, 71, 73, 125

JEDI Juno Energetic Particle Detector Instrument 65

JPL Jet Propulsion Laboratory 72

JUICE JUpiter ICy moons Explorer 65

LED Light Emitting Diode 224, 231, 235, 267

LEO Low Earth Orbit 7, 8, 11, 47, 59, 70, 73, 116, 120, 140, 173, 279, 283

LOFT Large Obervatory For X-ray Timing 15, 154, 156

LUCID Langton Ultimate Cosmic ray Intensity Detector 125, 127

MarCO Mars Cube One 72

MCP micro-channel plate 8, 11, 17–19, 79, 84, 119, 122, 205, 223–229, 231, 234–236, 239, 243–247, 251, 253, 255, 266–272, 274, 292

MEPED Medium Energy Proton and Electron Detector 140

MESSENGER Mercury Surface, Space Environment Geochemistry and Ranging 64

MHD magnetohydrodynamics 40, 41

MMS Magnetospheric Multiscale 82

MOSS Miniature Optimized Smart Sensor 89

MPPE Mercury Plasma Particle Experiment 64

MSSL Mullard Space Science Laboratory 33, 34, 87, 119, 135, 208, 210, 213, 214, 217, 223, 224, 232, 235, 275

Mulassis Multi-Layered Shielding Simulation Software 132, 158

MURA modified uniformly redundant array 147, 148, 152, 153

MuREM Micro Radiation Environment Monitor 85, 86, 91, 126

NASA National Aeronautics and Space Administration 13, 47, 48, 61, 89, 125

NIST National Institute for Standards and Technology 101–104

NOAA National Oceanic and Atmospheric Administration 140

OMC Optical Monitoring Camera 212

PACIDERM Plasma Automated Chamber Integrated Data Extraction Routine Miscellany 276

PBA Perfect Binary Array 16, 17, 147, 168, 169, 171, 188, 190, 191, 201, 203, 207, 213, 214, 237, 239, 258, 288

PC Personal Computer 19, 212, 222, 231, 276, 280, 281

PCB Printed Circuit Board 19, 218, 222, 281, 282, 289

PCFOV partially-coded field of view 155, 199, 201, 204

PEACE Plasma Electron And Current Experiment 82, 232

PMT photomultiplier tube 79, 80

PoleCATS Polar test of the Conceptual And Tiny Spectrometer 280

PSF Point Spread Function 16, 148–151, 171, 180–182, 253, 255, 256, 271

PSTAR STopping power And Range tables for protons 171, 180

PSU Power Supply Unit 280

RADFET radiation field effect transistor 85, 126

REPT Relativistic Electron Proton Telescope 86

REPTile Relativistic Electron Proton Telescope integrated little experiment 14, 86, 87, 91, 140

RHESSI Reuven Ramaty High Energy Solar Spectroscopic Imager 54

RM Radiation Monitor 85, 86, 91, 126

RXTE Rossi X-ray Timing Explorer 53, 152

SAA South Atlantic Anomaly 13, 53, 54, 61, 120, 173, 279, 288, 290

SEP Solar Energetic Particle 40, 46, 55

SHV Safe High Voltage 225

SLS Space Launch System 72

SNR signal-to-noise ratio 122, 145, 153, 184, 188, 264

SOHO SOlar and Heliospheric Observatory 80

SPECT single-photon emission computed tomography 142

SPENVIS Space Environment Information System 115

SPI Spectrometer on INTErnational Gamma-Ray Astrophysics Laboratory (INTEGRAL) 154

SREM Standard Radiation Environment Monitor 14, 84, 85, 91, 126, 127

SSTL Surrey Satellite Technology Limited 85, 86, 91, 126

STEIN SupraThermal Electrons, Ions and Neutrals 14, 89–91, 110, 124

STEREO Solar-Terrestrial Relations Observatory 45, 89

SWE Solar Wind Experiment 45

SWEPAM Solar Wind Electron, Proton, and Alpha Monitor 45

SXI Solar X-ray Imager 209

TDS-1 TechDemoSat-1 85, 87, 91

THEMIS Time History of Events and Macroscale Interactions during Substorms 67, 68

TID Total Ionising Dose 59

TIMAX Telescópio IMAgeador de raios-X 153

TRL Technology Readiness Level 290, 297

URA Uniformly Redundant Array 15, 25, 139, 145–149, 151–154, 163, 168–171, 175, 188, 213, 296

USB Universal Serial Bus 125

UV ultraviolet 224

WFM Wide Field Monitor 15, 154, 156

XML Extensible Markup Language 159

ZIF zero insertion force 220

Introduction

The space weather environment varies through time, and at each solar cycle large-scale systems on Earth and in orbit have become more complex and vulnerable to the Sun's effects. The field of space weather research is focused on understanding, predicting and monitoring the behaviour of the magnetosphere to prepare for these events.

This thesis is concerned with the investigation and proof of concept of a novel instrument design for future small scale space weather mission possibilities. Using very simple and low-resource components, which can take advantage of current and future developments in particle detection, computing power and telemetry, this instrument concept offers additional imaging capabilities compared to existing miniaturised detectors.

The coded aperture technique explored in this work, used with a position-sensitive silicon detector, has been simulated and tested in the lab, and demonstrated to provide a larger geometric factor and higher angular resolution than other instruments of comparable size and mass.

Space weather

The dynamic processes below and on the surface of the Sun produce large-scale events which penetrate through the solar system. The interaction of these events with Earth and its magnetosphere causes further effects in the particle and field

environments in orbit and on the ground.

These effects cause ever-increasing dangers for Earth and orbiting and ground-based infrastructure. Modern space-based systems including Global Navigation Satellite Systems (GNSS) and other communications satellites are particularly vulnerable to disturbances in the ionosphere, and all satellites are sensitive to high-energy radiation which causes electronic upsets. Human spaceflight is also affected, especially during Extravehicular Activities (EVAs). Any increase in radiation dose received is unhealthy and potentially deadly to astronauts. On Earth, pipelines and power supply networks create effectively a large antenna in which magnetic fields may induce unwanted currents.

Systems are now designed with some degree of robustness or self-preservation in mind. With knowledge and predictions of changes in the likelihood of space weather events over time, the effect of these dangers can be dealt with in the most cost-effective way. With specific knowledge of imminent events, further harm reduction techniques can be used. For example, flights can be rerouted away from polar regions to reduce communication and navigation issues, and satellites can shut off their most sensitive systems.

Therefore, missions both for research to more fully understand the processes involved in Sun-Earth interactions, and for monitoring events as they unfold in space before they reach the Earth are of great value.

Space weather instrumentation

Since the discovery of the radiation belts from measurements taken by the first artificial satellites, further missions have been planned and flown to understand them and their effects within the magnetosphere. Today there are many flying and planned missions and instruments both to research and to monitor space weather effects remotely and in-situ.

Many current instruments rely on large spacecraft and launch vehicles to reach

orbit or beyond, and such missions are only possible with a long development time and significant funding.

More recent developments in technologies involved in this instrumentation have allowed the development of a new family of capable small-scale, low-resource instruments. These instruments may even fit on a CubeSat platform, the relatively recent form factor of satellites. This type of satellite has a short development timeframe and the ability to find a launch opportunity quickly, as well as increased possibility for multipoint measurements, especially of new environments.

A proposed instrument with high geometric factor

This project aims to take advantage of the recent developments in hardware, production techniques, electronics and computational power as well as heritage from previously flown instruments of different types. This can then produce a novel particle detector for a small satellite platform which can work with recent related developments in instruments and detectors of fields and particles for space weather research.

Based on previous work at the Mullard Space Science Laboratory (MSSL) into high angular resolution particle detection for space weather research, a complementary device was proposed, using a coded aperture mask in conjunction with a position-sensitive detector. This trades the ability to identify the incoming direction of an individual particle for the ability to provide information on the particle environment in low density or low flux environments.

The coded aperture concept is based on similar, but far larger, instruments that have been used for decades in space-based X-ray and gamma-ray astronomy.

The miniaturisation of such instruments relies on the availability of small detectors, small and low-power electronics, and any additional support for these systems such as cooling. Although this limits the possibilities, recent developments in technology have made several types of such components available.

Simulations

To investigate this proposed design, a full computational simulation of such an instrument was constructed using Geant4, a toolkit developed for the simulation of the transport of particles in detectors and other materials.

These simulations were run for many types of energy and angular distributions for both electrons and protons and for several types of instrument geometry in order to better understand the design decisions.

Simulations were also used during the testing phase of the development to understand the results, and when considering a full flight instrument in a space environment.

Testing the concept

In order to demonstrate the validity of the calculations and simulations, a lab-based proof-of-concept demonstration instrument and test setup was assembled. Various trade-offs needed to be made, both because of the general limitations of a laboratory environment when attempting to simulate a space environment, and the specific limitations of what could be achieved in the time and with the facilities and hardware available at MSSL.

The concept was demonstrated in a vacuum chamber using a high-end charge coupled device (CCD), which was the best available position-sensitive particle detector, and a copper-tungsten pseudo-alloy coded aperture mask.

A number of electron source geometries were used in an attempt to demonstrate the ability of the system to recover the expected source distribution. Three sources were used with different energy spectra and angular distributions - two disc-shaped radioactive beta sources, samarium-151 and carbon-14, and a modified micro-channel plate (MCP)-based electron source with a grid creating an acceleration potential at the front.

Although the availability of sources limited the conclusions that could be drawn

from the experiments, the understanding of the CCD and the sources, especially the MCP source, was improved.

Conclusions

The simulations and the experimental data allowed several analyses to be performed. The behaviour of the CCD and the accompanying setup in response to temperature and radiation was quantified, and both the radioactive sources and the MCP-based source were further characterised. However, there were limits to what could be tested in the laboratory due to the available hardware.

The lab-based instrument allows improvements in geometric factor and angular resolution for energetic particle measurements. However, this construction was designed as a proof-of-concept and is not optimal for spaceflight, for reasons including the necessity of cooling the detector and the sensitivity to light. The detector pixel size is also unnecessarily small compared to the scale of the mask and the size of the point spread function of an energetic particle hit.

A more suitable instrument has been proposed and simulated in a space-like environment. This type of instrument would be able to analyse the energy and angular distributions of particle populations, but would remain vulnerable to situations where light would enter the instrument.

Design context

This proposed instrument would need further development to be optimised for use in a specific space environment.

The initial simulations and experiments demonstrate a flexible instrument, which in some environments would be a strong addition to the current available technology. This instrument, or one of its descendants, could be used together with other small-scale plasma instrumentation for both particles and fields.

This thesis is only the first step in the development process of this type on instrument. A sounding rocket or a CubeSat would be the ideal platform to demonstrate the concept in a relevant environment.

Chapter 1

Space plasma and particle environments

The environment in near-Earth space is of great interest to scientists and providers of space-enabled services such as telecommunications, navigation and climate and weather forecasting. The results produced by data-gathering missions can be used to further our understanding of the heliosphere and magnetosphere and are of critical importance to many technologies, both satellite and ground-based. Since the beginning of the space age, *in-situ* particle instrumentation has worked alongside all other space systems both to learn more about the behaviour of the dynamic magnetosphere and to inform the design and testing of more reliable and accurate spacecraft and space systems in harsh and varying environments. As satellites and their instrumentation become more important to systems on Earth, and network infrastructure on Earth becomes larger and more vulnerable, for example to induced currents, an understanding of the effects of the changing behaviour of solar particles and plasmas on our environment becomes more critical. As a result, more and more efforts have been made to further this understanding with a variety of science missions, including those in low-Earth orbit and throughout the magnetosphere.

In more recent decades, the advances in miniaturisation of manufacturing and

computing power have allowed smaller instruments to be flown at lower cost, opening up new ways of gaining these data, for example swarm missions and additional low-resource instruments on commercial satellites.

1.1 Matter in space

Outer space is far from empty. The vast volumes between the Earth, Sun, planets and other bodies within the solar system are permeated with a range of particles and fields, the vast majority of which is in the form of plasma, a state of matter found in the hot and sparse conditions of the heliosphere. On Earth, plasmas are present in natural phenomena such as flames and lightning, and also in human-made objects such as fluorescent lights and tokamaks, but in space they are almost ubiquitous: energetic particles such as those produced by the Sun spread throughout the interplanetary environment both creating and interacting with electric and magnetic fields. It is the work of many scientists to piece together an understanding of the chains of events and interactions between processes below the surface of the Sun and effects at the ionosphere of the Earth and other planets; the data from multiple instruments of many types is required to support this research.

On larger scales under the highly dynamic influence of the Sun, the behaviour of plasmas, particles and fields is far from completely understood, but they follow well-understood patterns on a microscopic scale, being guided by the Lorentz force, and following Maxwell's equations. The effects of magnetic fields on individual particles shown in this section are key for understanding many of the macroscopic plasma effects such as the radiation belts of Earth.

1.1.1 Particle behaviour in magnetic fields

A plasma is defined as a state where particles are sufficiently energetic that some or all of them have separated to some degree into electrons and charged ions; they are only able to exist at low densities where the recombination time of electrons and

ions is long. The behaviour of such a system is significantly different from matter in other states. Electrons and ions may behave as separate populations, for example being at different temperatures, or accelerating in different directions or at different speeds as a result of their different particle masses and charge signs.

Particles within a plasma will accelerate when gravitational or magnetic fields are present, including magnetic fields that have been created by the movement of the charged particles themselves, as a straightforward effect of the Lorentz force

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} \tag{1.1}$$

where the force $\mathbf{F}$ is proportional to the cross product of the particle velocity and the magnetic field vectors $\mathbf{v}$ and $\mathbf{B}$

For example, a charged particle moving in a uniform magnetic field follows a helical path: the Lorentz force, being perpendicular to the sideways motion, causes a circular motion, while the component of the motion parallel to the magnetic field is neither accelerated nor decelerated. In this way, charged particles are usually constrained to move along or around magnetic field lines as shown in figure 1.1. This happens in flux ropes above the Sun's surface or radiation belts in the Earth's magnetosphere; negatively-charged particles gyrate anticlockwise around the magnetic field as viewed against the magnetic field direction, while positively-charged particles gyrate clockwise.

The properties of this simple gyration can be easily described for a non-relativistic case: the gryofrequency is

$$\Omega_c = \frac{|q|\, B}{m} \tag{1.2}$$

where the frequency is proportional to the magnitude of the charge on the particle and the magnetic field strength, and inversely proportional to the mass of the particle. Therefore, the frequency of gyration of an electron is approximately 1800 times higher than that of a proton in the same magnetic field.

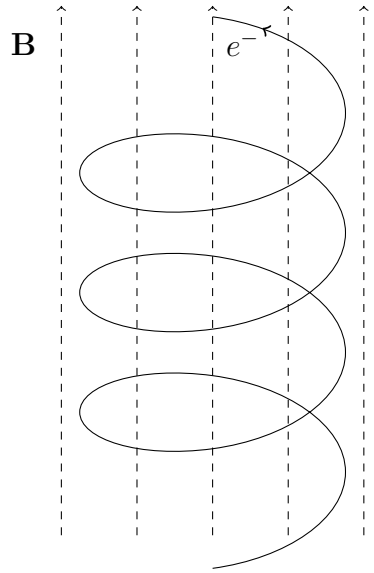


Figure 1.1: Path of an electron in a uniform and constant magnetic field

From this, it can be shown that the gyroradius, or Larmor radius, is

$$r_c = \frac{v_\perp}{\Omega_c} \tag{1.3}$$

where $v_\perp$ is the component of the particle's velocity perpendicular to the magnetic field. This gyration is independent of the particle's movement parallel to the field lines.

In addition, a non-uniform magnetic field may cause a drift in the position of the central axis in the gyration; an example of this is the drift of radiation belt particles around the Earth described in section 1.3.3.

1.1.2 Magnetic pressure and particle acceleration

Magnetic pressure is a concept used to describe the pressure-like force on a charged particle pushing it away from areas of high magnetic field density. Mathematically,

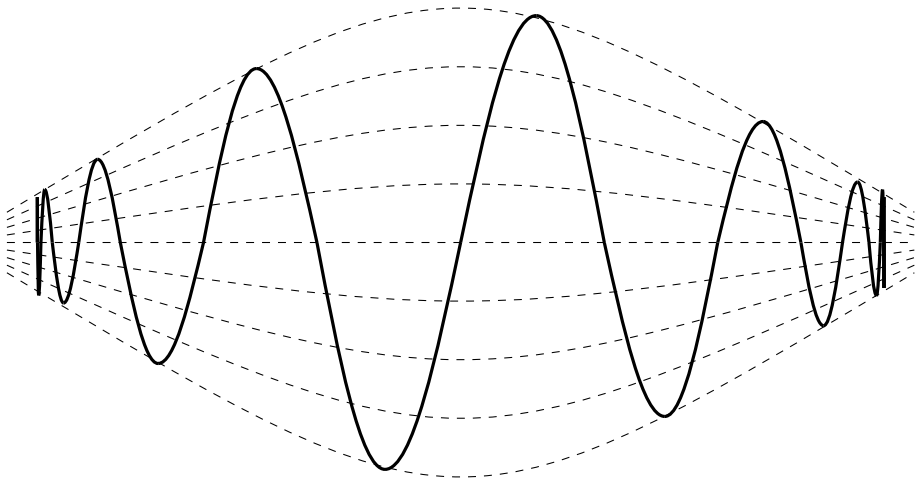


Figure 1.2: Trajectory of a charged particle trapped between magnetic field lines, shown in two dimensions, after Gombosi [1998]

the magnetic moment of the particle is conserved, and the force increases with the component of the particle's velocity which is perpendicular to the magnetic field lines. The force produced by this pressure can be found by taking the cross product of the Lorentz force as given in equation 1.1 with the magnetic field strength vector to give

$$\mathbf{F} = -\nabla \left(\frac{B^2}{2\mu_0} \right). \tag{1.4}$$

Figure 1.2 shows a schematic magnetic trap, a concept which has been used to confine plasma in large-scale ground-based fusion experiments, but is also representative of certain structures in the magnetosphere and heliosphere. The increasing magnetic field at either end of the trap reduces and then reverses the component of the particle motion parallel to the magnetic field lines, changing their pitch angles relative to those lines, and keeping the majority of particles within the confines of

the trap. However, particles travelling in directions close to parallel to the magnetic field are relatively unaffected by the magnetic pressure and those at the most extreme angles are able to continue through the ends of the magnetic trap.

In this situation, the conservation of magnetic moment can be expressed as

$$\frac{\sin^2\theta}{B} = const \tag{1.5}$$

where θ is the pitch angle of the particle with respect to the magnetic field lines, and B is the local magnetic field strength.

The particles that are capable of escaping the trap may be identified by the use of this equation; particles that will not reach a pitch angle of 90° by the maximum magnetic field strength will still have a parallel velocity component out of the trap, which can be expressed as

$$\sin^2\theta_{min} = \frac{B}{B_{max}}. \tag{1.6}$$

Particles with a greater angle will reach a mirror point and remain within the trap.

Additionally, strong temporal changes in the plasma environment such as those formed at shocks associated with coronal mass ejections (CMEs) in the heliosphere can cause strong and significant accelerations, in some cases creating a population of highly energetic charged particles such as Solar Energetic Particles (SEPs).

1.1.3 Magnetohydrodynamics

The field of space plasma physics relies heavily on the applications of such interactions integrated into a framework known as magnetohydrodynamics (MHD), which can then be used to model the environments and situations described later in this chapter. MHD combines Maxwell's equations and a modification of the Navier-Stokes equations of fluid dynamics into one system describing possible behaviours of particles and fields of plasmas in the solar system; its first use by name was probably by Alfvén [1945], but modern techniques use both analytical and computational

methods to attempt to reconcile our physical understanding of plasma processes with the observed behaviour. An overview of the techniques and applications of MHD are given in Priest and Forbes [2000] and others.

1.2 The heliospheric environment

The Sun's dynamic surface drives the motions and processes of interplanetary and planetary environments.

The Sun is continuously losing mass to the heliosphere, as particles in the solar atmosphere are accelerated to the energies needed to escape the gravitational pull. This particle loss is approximately $2 \times 10^{-10}\,\mathrm{g\,cm^{-2}\,s^{-1}}$ in coronal holes and up to $4 \times 10^{-11}\,\mathrm{g\,cm^{-2}\,s^{-1}}$ in active regions, according to Aschwanden [2006].

The Sun produces this solar wind from its entire surface, though at higher latitudes the wind is usually more energetic. Solar winds vary between $400\,\mathrm{km\,s^{-1}}$ - the slow solar wind, which has a composition similar to the solar corona - and $750\,\mathrm{km\,s^{-1}}$- the fast solar wind, which has a composition similar to the photosphere. Results of the measurements from the Ulysses mission, which studied the physics of the heliosphere, are shown in figure 1.3. This continuously-produced solar wind travels through the heliosphere, with the energy spectrum of the particles forming a Kappa distribution.

This environment is of great interest for space weather predictions. All plasma behaviour in the Earth's magnetosphere - the region influenced by the Earth's magnetic field - is linked to the source and behaviour of this energetic plasma. Research into the processes joining the happenings on the solar surface to the effects on Earth, and into measurements upstream in the solar wind - both in-situ and remote - can improve the speed and accuracy of space weather predictions.

The rotation of the outer layers of the Sun, the speed of which varies by latitude, gives rise to the Parker spiral which was described by Parker [1958] and is shown in figure 1.4. Magnetic field lines, anchored in a feature of the solar surface, are drawn

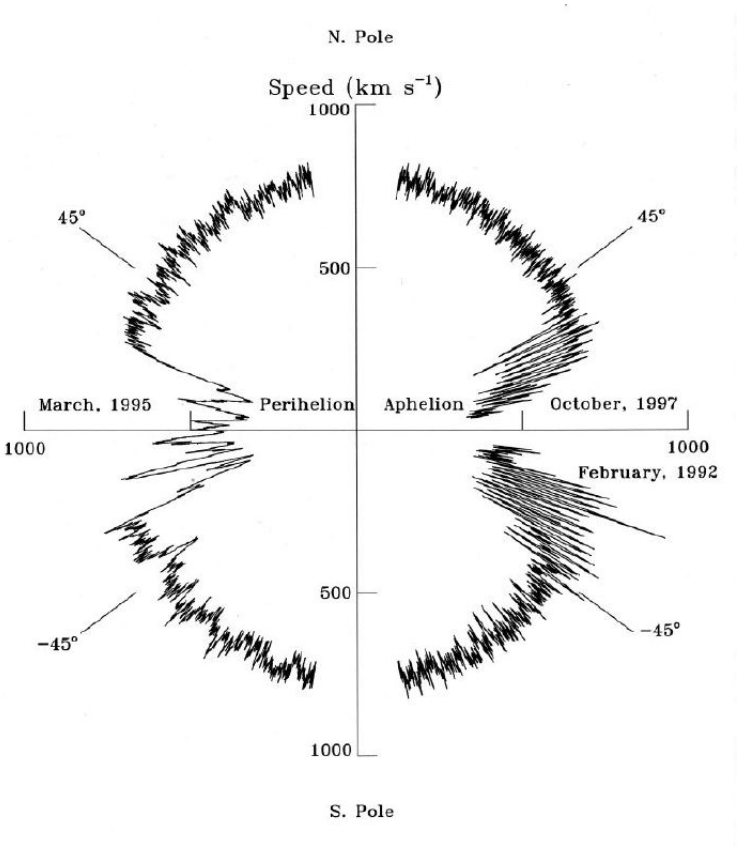


Figure 1.3: Solar wind speed against solar latitude as measured by the Ulysses spacecraft from McComas et al. [1998]

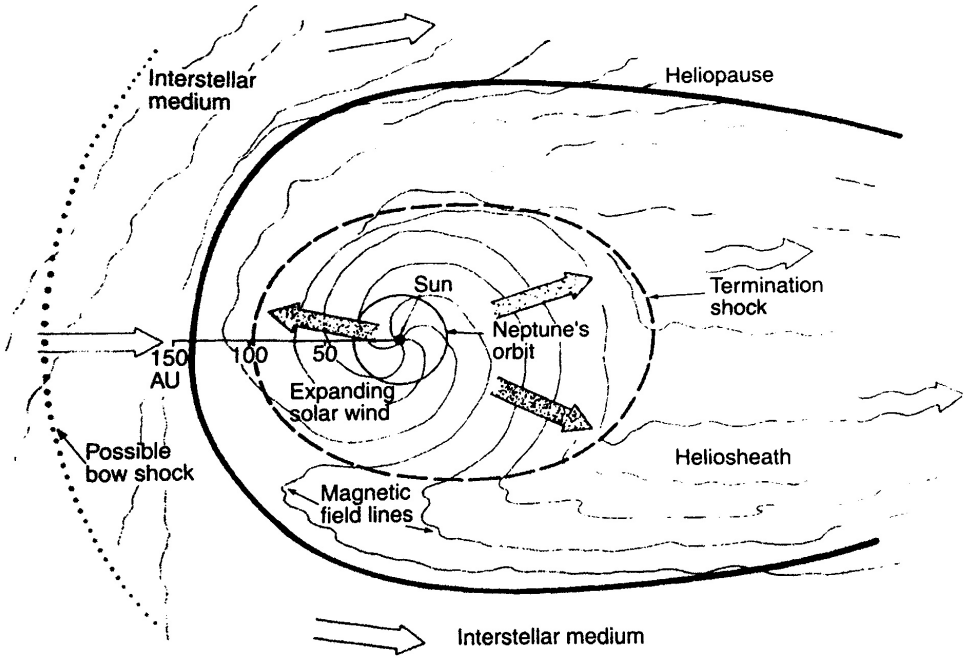


Figure 1.4: Schematic diagram of the heliosphere as seen from the ecliptic north pole from Gombosi [1998]

out into a spiral shape through the heliosphere, and lower-energy charged particles are constrained to move along these lines. It can be therefore be expected that from the point of view of a spacecraft, the flow of solar wind particles will be away from the Sun, but also directed from a position to the west of the solar disk.

1.2.1 Co-rotating Interaction Regions

A co-rotating interaction region (CIR) is a consequence of the variability of the solar wind production at the Sun's surface and the rotation of the Sun. Areas of the solar surface which produce higher-speed and slower-speed streams create spiral-shaped regions of fast and slow plasma throughout the heliosphere. These fast and slow

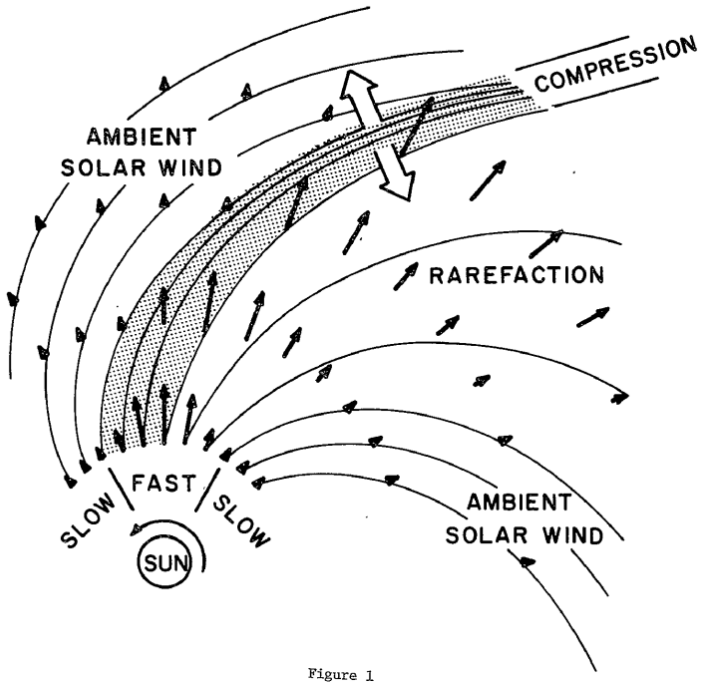


Figure 1.5: Geometry of the interaction between fast and slow solar wind from Pizzo [1978]

flowing areas interact to produce high-density areas which, at distances beyond the Earth's orbit may become significant enough to produce shocks around the zone of compressed plasma. The cause of CIRs is shown diagrammatically in figure 1.5, with the CIR shown as the compressed and rarefied areas at the interaction.

1.2.2 Coronal Mass Ejections

Events on the solar surface can frequently but sporadically lead to a massive outburst of plasma and magnetic fields known as a CME, which can have a mass of 10^{14} g to 10^{16} g Aschwanden [2006]. The processes leading to the production of CMEs are not fully understood, and predicting geo-effective solar events with the maximum

warning time is an important aim for future missions.

In particular, the ambiguity between CMEs in the directions directly toward or away from Earth is not always clear. These CMEs are far easier to observe and distinguish either by using a vantage point to the side of the Earth-Sun line, as the Solar-Terrestrial Relations Observatory (STEREO) spacecraft do, or using in-situ measurements when it is possible to locate a spacecraft between the Sun and the Earth. When a CME does interact with the Earth's magnetosphere, the orientation of the magnetic field in relation to the direction of the Earth's magnetosphere is crucial to the geo-effectiveness of the interaction due to the energy released in reconnection. This analysis of field configuration is very difficult to measure remotely and can be far more accurately calculated from *in-situ* measurements.

1.2.3 Energetic particle acceleration in the heliosphere

Fields, shocks and waves can cause certain populations of particles in the heliosphere to be accelerated beyond the usual energies of solar wind particles.

Electron strahl

Strahl is the sharply magnetic field-aligned suprathermal electron population within the solar wind, which exists alongside the lower-energy 'core' and 'halo' populations which are not field-aligned, as described by Pierrard et al. [2001]. Although the energy range of strahl is not clearly defined, these strahls have been identified from measurements of electrons between around 50 eV and 1370 eV, although at higher energies in this range, the number of particles is lower and statistically, the strahl is harder to analyse.

Measurements of the angular and energy distribution of strahl electrons have been gathered using the Solar Wind Experiment (SWE) package on the Wind spacecraft and the Solar Wind Electron, Proton, and Alpha Monitor (SWEPAM) package on Advanced Composition Explorer (ACE) as well as higher-latitude measurements

taken from Ulysses. Several analyses of these data have been done, all of which have found large variation in the angular width of the flux distribution.

Anderson et al. [2012] have determined strahl widths to be from 5° to 90° but concluded that no typical angular width for strahl for a certain radial distance from the Sun could be determined. The width is determined by the competition between particle scattering widening the beam and magnetic focusing, which narrows the beam. Narrower strahl flows below about 20° wide are often associated with high-speed streams, but strahl width may either decrease or increase with energy.

Hammond et al. [1996] have investigated the change in strahl behaviour for multiple solar latitudes over the lifetime of the Ulysses spacecraft, where the width was measured to be broader than expected. The average strahl width adjusted to be equivalent to a solar distance of 1 AU was found to be 49° for 77 eV electrons, which was the energy for which there was the broadest distribution. The strahl width was found to increase between 1.3 AU and 2.3 AU after which it was approximately constant. Scattering processes are responsible for the broadening of strahl, which would otherwise be more highly collimated.

Solar energetic particles

SEPs are those populations accelerated to between a few keV and several GeV from flare sites and shock waves in the solar corona. Schwenn [2005] has produced a review of the current understanding of the production and behaviour of SEPs.

Energetic particles can be released directly from solar flares, in which case they will produce a very sudden event, or from acceleration over CME shocks, in which case a more continuous event is produced. The most energetic particles can reach Earth from the solar surface in a few minutes, and are a particular danger to space systems.

SEPs usually contain a distribution of more massive ions, which are representative of the ion distribution in the solar atmosphere.

1.3 The Low Earth Orbit (LEO) spacecraft environment

LEO is generally defined as a low-eccentricity orbit below an altitude of 2000 km. A satellite's inclination may allow it to focus on the equatorial, auroral or polar regions.

The region covering the upper thermosphere and inner exosphere up to approximately 400 km, which is the environment of a large number of smaller satellites and the majority of CubeSats, is highly stratified, with widely varying characteristics across altitude and latitude.

Figure 1.6 by NASA shows some of the known behaviours of the ionosphere where the Earth's atmosphere is partially or fully ionised, which will influence the environment around these missions. The most significant and interesting features of the plasma environments in space are given in this section.

Dose rates and radiation profiles measured by instruments on previous spacecraft with similarly low orbits, such as the International Space Station (ISS), can be used as an indication of the expected radiation environment at the similar altitude and latitude sections of LEOs.

1.3.1 Temperature and composition profiles

The temperature varies between around 190 K and 1000 K in the altitude range of LEO, from roughly 100 km to 2000 km, as shown in figure 1.7. Above the mesopause (the temperature minimum between the mesosphere and the thermosphere at around 85 km to 100 km), the temperature rises rapidly and close to linearly until levelling off at around 1000 K.

With increased altitude, there is a higher degree of ionisation, where space plasma and radiation interact with the atmosphere. In addition, the composition changes. Below the turbopause (at around 110 km), turbulent currents keep the proportions of atmospheric constituents approximately constant, but above this altitude, the

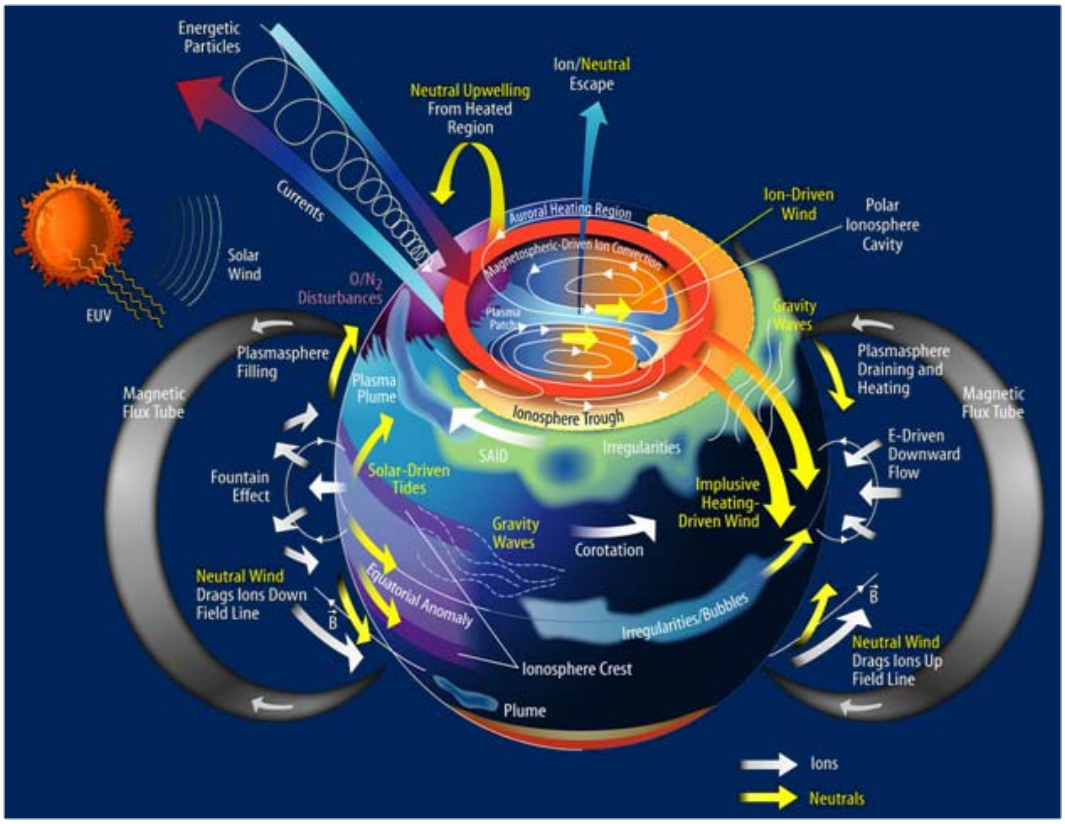


Figure 1.6: National Aeronautics and Space Administration (NASA) diagram of the most significant ionospheric currents

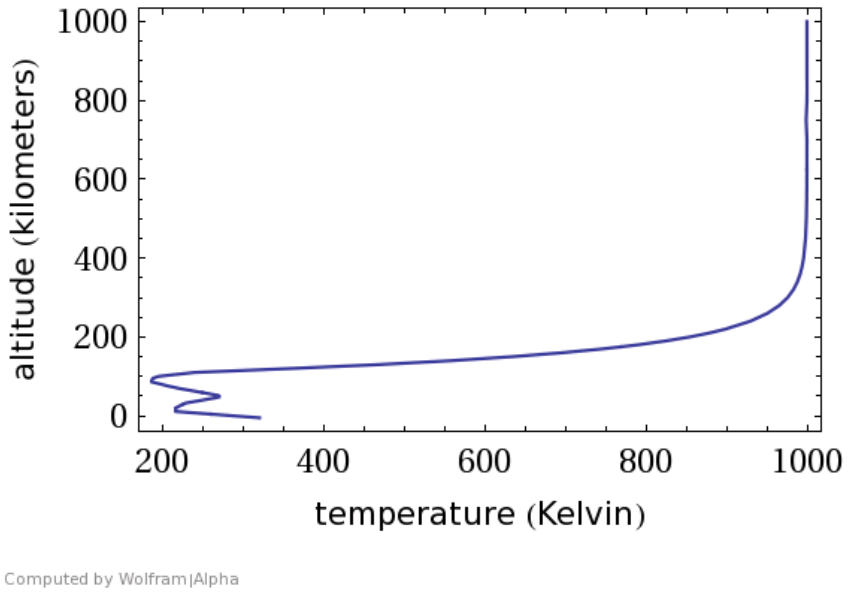


Figure 1.7: Average values for atmospheric temperature against altitude given by WolframAlpha [2016]

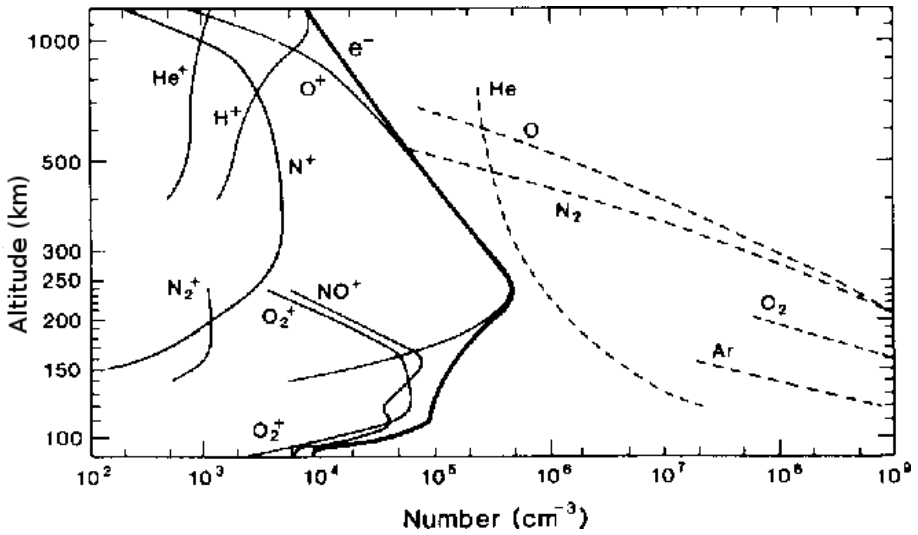


Figure 1.8: Average values for atmospheric neutral and ion composition against altitude from Johnson [1969]

ions and atoms form layers, the distribution of which depends on the local time of day and the latitude - figure 1.8 shows an average atmospheric composition based on mass spectrometer measurements.

1.3.2 Atmospheric gravity and planetary waves and tides

Waves and tides in the atmosphere can occur on several preferred scales, either localised or global.

Thermospheric tides become more important at greater heights, described in detail in, for example, Hargreaves [1995]. Planetary waves have wavelengths of the order of magnitude of the radius of the Earth. Tides have frequencies which are harmonics of a day - usually 12 or 24 hours.

Acoustic Gravitational Waves (AGWs) are often caused by localised atmospheric perturbations, for example by thunderstorms, flows over irregular mountains or auroral activity. It has been found by Skorokhod and Lizunov [2012] from satellite

measurements that there is a regular component of the wave distribution, and the wave packets are more likely to occur near the poles than the equator. Kryuchkov and Fedorenko [2012] have demonstrated mathematically the optimal properties of these waves for energy transport; at the upper regions of the polar atmosphere, the waves observed by the Dynamics Explorer probes were seen to have these frequencies and wavelengths. The waves propagate through the atmosphere because the disturbance causes an imbalance between the forces of weight and pressure of the atmosphere.

These small-scale disturbances are the type of event a very low-altitude constellation mission would be in a very strong position to investigate.

1.3.3 The radiation belts

The radiation belts are caused by the shape of the magnetic field lines of the Earth; electrons and protons in particular are confined to gyrate along the field lines. Where they approach the poles, magnetic pressure, as described in section 1.1, reverses their direction parallel to the field, trapping these particles within the belt.

Particles will spiral around magnetic field lines and be repelled by the increasing magnetic pressure when they near the poles to reflect back and forth around the radiation belts; these are known as trapped particles and gyrate back and forth between the poles along the radiation belts. Particles which are most closely aligned with the direction of the magnetic field lines however may escape downwards into the atmosphere in the polar regions. This removal of particles leaves a population of electrons, protons, and to a lesser extent heavier ions, which bounce between north and south poles gyrating in a motion which depends on the sign of their charge, at a speed which depends on their energy. In addition to this relatively fast motion, the circular motion in a non-uniform magnetic field causes a sideways drift of particles; positive ions drift clockwise, and electrons drift anti-clockwise, creating a small current around the Earth.

In-situ observations from the very early years of the space age were shown by Van

Allen [1958] and Van Allen [1959] to indicate that the Earth's Van Allen radiation belts could be separated into an inner zone dominated by high-energy protons and an outer zone dominated by high-energy electrons. These belts continued to be mapped out more fully by further missions in the years following these studies.

Later studies, including those by Gussenhoven et al. [1996] and Heynderickx [2002] based on data from the Combined Release and Radiation Effects Satellite (CRRES) showed that electrons of energies up to around 1 MeV often populate both zones, with a region largely devoid of particles between them. The two-belt structure was explained as arising from strong electron interactions with the cooler denser plasma just inside the plasmapause boundary. However, recent data analysed by Baker et al. [2013] has shown that a third belt is present at times, when solar activity causes changes which allow two belt structures to take the place of the outer belt.

Even more recent measurements from the Van Allen Probes were shown by Baker et al. [2014] to demonstrate the presence of a further structure within the belts: a very sharp plasma boundary at the inner edge of the outer belts which prevents high-energy (greater than around 5 MeV) electrons from penetrating.

As a result of these structures, the particle flux environment depends heavily on the position in the radiation belts, while the position and exact behaviour of the belts and of the particles within the belts is highly variable and especially dependant on solar influences.

1.3.4 South Atlantic Anomaly

Over the South Atlantic Ocean, the Earth's magnetic field is significantly weaker than the field at similar altitudes at other points on the Earth. The radiation belts are also at an unusually low altitude. Together, these form an area where an abnormally large number of energetic particles are allowed to penetrate into the Earth's atmosphere, where they cause ionisation of atmospheric constituents at lower altitudes, with a peak in ionisation at 80 km as described by Abel and Thorne

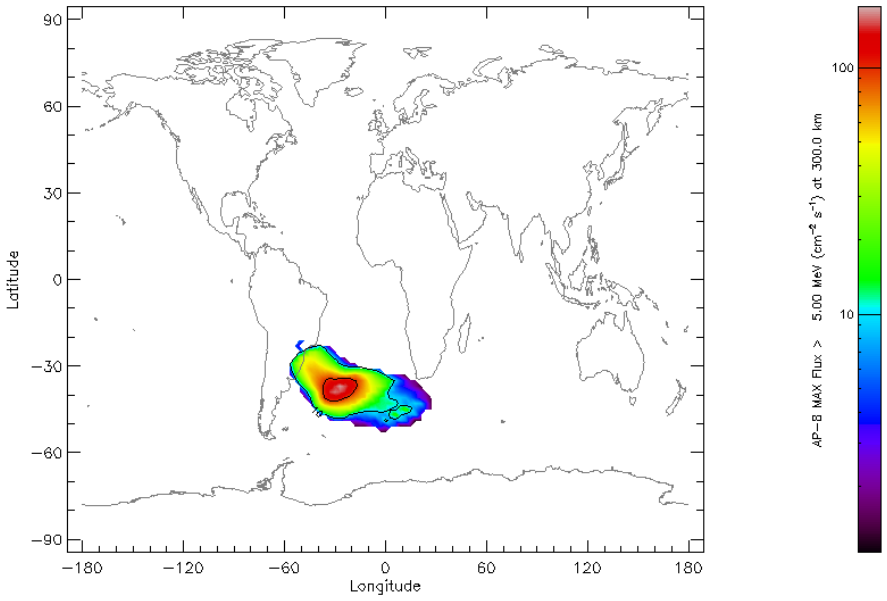


Figure 1.9: Global flux of protons above 5 MeV at 300 km altitude at solar maximum. Data are from the AP-8 atmospheric model, via SPENVIS (Available at www.spenvis.oma.be)

[1999]. These particles are largely electrons and protons, with a smaller fraction of ions. The energetic protons and magnetic fields in this region are particularly problematic for satellite hardware; because of the frequency of problems and errors caused within the electronics and processors as described in section 1.3.8, some less-protected satellite systems will routinely turn off while passing through the South Atlantic Anomaly (SAA). Figure 1.9 shows the extent of the anomaly at an altitude of 300 km.

The SAA has several properties worthy of investigation. The shape and movement of the anomaly have been studied in the past, for example by Fürst et al. [2009], who have used measurements from the Rossi X-ray Timing Explorer (RXTE). It

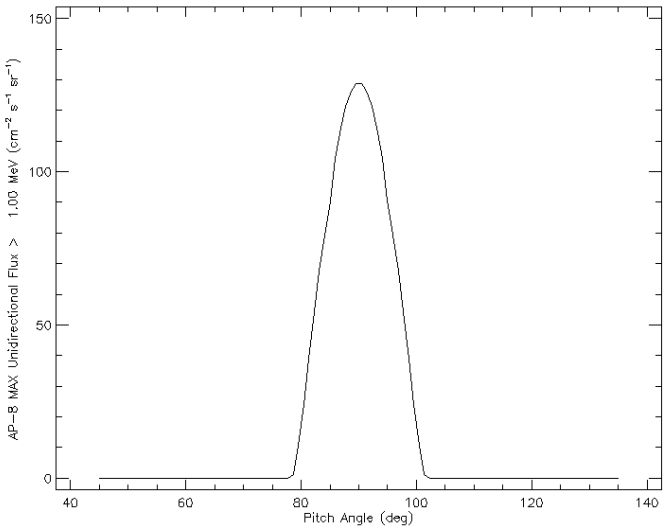


Figure 1.10: Angular distribution of the trapped proton population in the SAA at 300 km altitude. Data are from the AP-8 atmospheric model, via SPENVIS (Available at www.spenvis.oma.be)

was found in this article that the distribution in space was best-fitted by a Weibull distribution varying in size and position, while Hell and Bamberg [2010] fitted a Gumbel distribution to the central area of the anomaly estimated from Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) data. The centre of the anomaly drifts west over time at a current average speed of about 0.25° each year, but with a variation of up to around 1°, as measured by Fürst.

Trapped particles (electrons, protons and ions) at these altitudes are usually restricted to move within a very narrow band of horizontal angles, as shown in figure 1.10; distributions such as this, found in the auroral regions and SAA, are known as “pancake distributions”.

1.3.5 Auroral regions

In the auroral regions (above 70° magnetic latitude), the pattern of activity and behaviour of the particles and currents in the ionosphere is a projection of the magnetosphere. The ionised particles entering the ionosphere from the radiation belts and magnetosphere can give an indication of their origins in the magnetic field. The ionosphere is consequently heated, both by auroral particle precipitation and resistive heating by currents driven by the magnetospheric dynamo, in addition to the solar heating otherwise experienced. Current results, for example from the Cluster mission, have provided some insight into the behaviour of such particles travelling along the field lines in other parts of the magnetosphere.

Models of the flux of electrons in the polar cusps were described by Banks et al. [1974] and later by Luhmann [1976], which include the attenuation of lower-energy electrons at lower altitudes. Figure 1.11 shows the energy distribution of the electron flux in the auroral zone, based on Banks' and Luhmann's models.

These influxes of SEPs and radiation into the cusps at the auroral oval can cause effects right down to the Earth's surface; nitric oxide ions are created which can destroy ozone molecules, and the ionospheric disturbances can disrupt the propagation of radio waves. Small-scale or short-lived structures in the ionosphere often cause disruption to the reliability of satellite positioning systems, and currently large-scale simple models are used to account for these effects. Data on shorter distance and time frames could inform improvements in these models and add reliability to GNSS measurements.

Measurements of electron densities in and around auroral arcs have been measured on multiple sounding rocket campaigns, for example by Venkatarangan and McEwen [1979] and Ogasawara et al. [2006], which were capable of high-cadence in-situ measurements of electrons up to the order of tens of eV. Inclined orbital missions can spend a large proportion of their orbit in auroral regions.

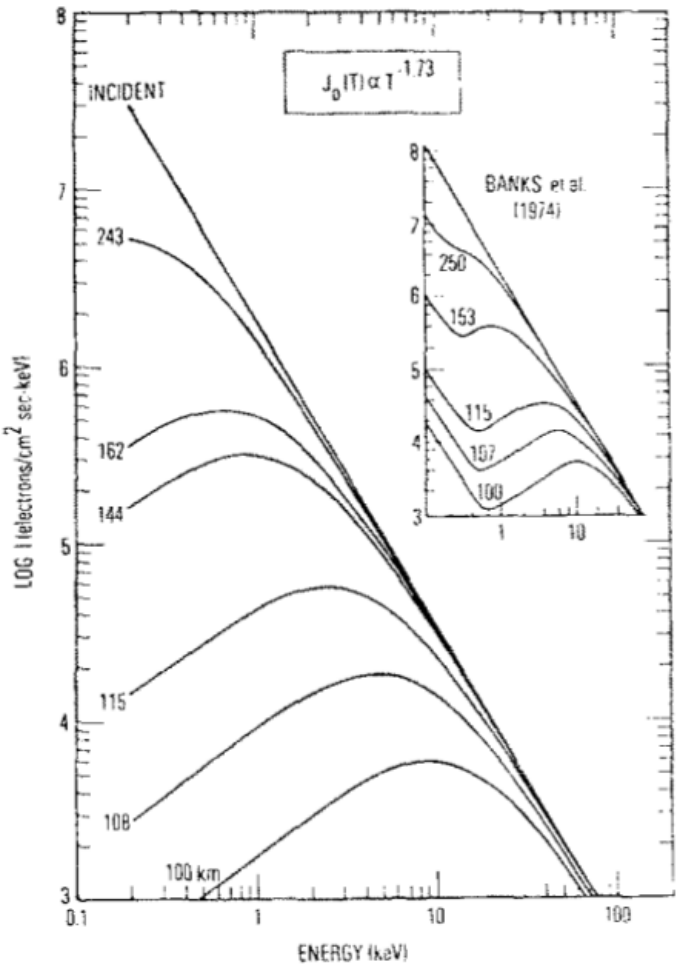


Figure 1.11: Hemispherical flux spectra downwards at several altitudes from Luhmann [1976]. Inset is the equivalent calculations by Banks et al. [1974], which takes into account secondary electrons.

1.3.6 E- and F-region dynamos and currents

The Equatorial Electrojet (EEJ) is a band of eastwards-directed current on the dayside of the Earth along the equator in the E-region of the ionosphere between around 90 km to 150 km altitude.

There also exist auroral electrojets circling the north and south poles. The movement of such particles creates perturbations in the magnetic fields right through the atmosphere, and can be detected using ground-based magnetometers.

In-situ measurements of currents in the E-region of the ionosphere have not been taken from orbit since those from the highly-elliptical orbit satellites between the 1960s and the early 1980s as noted by, for example, Skorokhod and Lizunov [2012].

1.3.7 Space weather

By far the largest source of energy to the atmosphere is the Sun, and this energy is transmitted in the form of both radiation and particles. As a particle source, the Sun is a highly dynamic and variable influence. Events created far below the solar surface can cause massive effects right through the heliosphere. The production of the solar plasma and events within it were described more fully in section 1.2.

The Earth's atmosphere is protected from the majority of the particle radiation by the fields in the magnetosphere, but the environment in and under the magnetosphere is highly changeable, and depends heavily on solar activity. The science of the understanding and prediction of the magnetosphere, ionosphere and atmosphere in reaction to solar activity is known as the field of space weather.

In general, the shape of the Earth's magnetic field and magnetosphere is pulled from its dipole shape by the incoming magnetic fields from the Sun, which wrap around the magnetosphere and pull it outwards. This shape is shown in figure 1.12.

The effects of solar influences can change the behaviour of Earth's ionosphere on very short timescales; an event's major effects will typically last for the order of a day, and changes in the aurora can be observed to move rapidly, sometimes on a

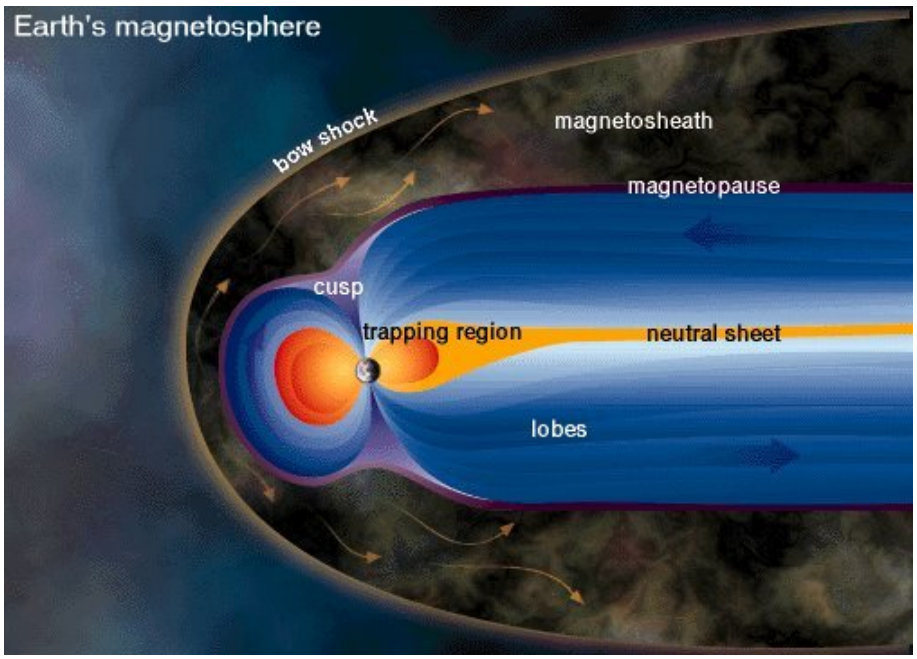


Figure 1.12: The shape of the Earth's magnetosphere distorted by pressure from the Sun's radiation (image from NASA)

timescale of seconds.

Solar behaviour and consequently solar effects on Earth are heavily correlated with a well-established eleven-year cycle, first described by Schwabe [1844]. Sunspots and solar events occur far more often during solar maximum than during solar minimum. In addition to higher particle fluxes and more changeable environments, it is also observed that Galactic Cosmic Ray (GCR) numbers are significantly lower during solar maximum. Figure 1.13 shows the difference in GCR fluxes at the lower end of the energy spectrum between solar minimum and maximum.

1.3.8 Radiation effects in LEO environments

At altitudes and coordinates affected by high fluxes of high-energy particles, especially in and around the radiation belts, both single energetic particles and strong magnetic fields can cause problems for electronics, processors and also astronauts.

An energetic proton can contain enough energy to flip a bit in the memory of unprotected electronics, causing an error in data or software or a latchup which can disrupt, damage or destroy software and hardware, while excessive Total Ionising Dose (TID) can also cause the degradation or failure of sensitive electronics or sensors. These issues are difficult to deal with and instruments designed to operate in the radiation belts will have to make either the processing hardware or software radiation-hard. The famous Carrington event in 1859 described by Carrington [1859] and Hodgson [1859] was an unusually severe solar event, and had by far the most severe effects on Earth in recorded history; such an event in the age of satellites and sensitive large-scale systems would undoubtedly cause serious consequences to infrastructure.

Table 1.1 lists the environmental hazards in harsh plasma environments; these are described by Vampola [2000] and others.

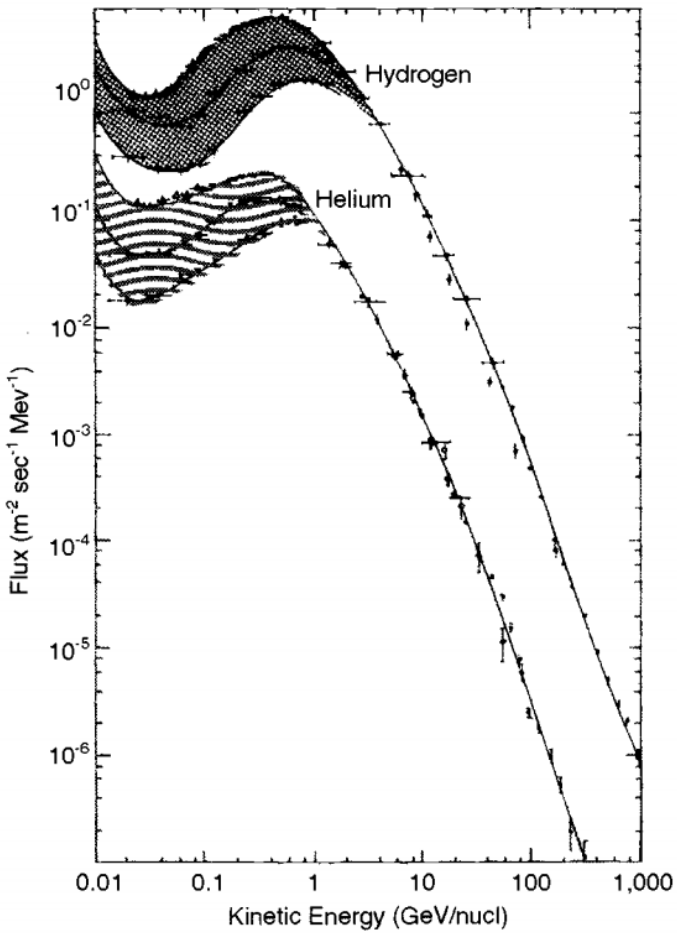


Figure 1.13: The GCR spectra at solar minimum and maximum, showing the differences at lower particle energies, from Gombosi [1998]

Table 1.1: Space particle hazards as quoted in Ginet and O'Brien [2009]

Environmental hazard	**Particle population**	**Natural variation**
Surface Charging	0.01 keV to 100 keV e^-	Minutes
Surface Dose	0.5 keV to 100 keV e^-, H^+, O^+	Minutes
Internal Charging	100 keV to 10 MeV e^-	Hours
Total Ionizing Dose	>100 keV H^+, e^-	Hours
Single Event Effects	>10 MeV/amu H^+, Heavy ions	Days
Displacement Damage	>10 MeV H^+, Secondary neutrons	Days
Nuclear Activation	>50 MeV H^+, Secondary neutrons	Weeks

1.3.9 Atmospheric and ionospheric models

A review of atmospheric and ionospheric models is given by Belehaki et al. [2009].

In 1965, NASA produced the first iteration of aerospace particle models (AE-1 and AP-1) for electrons and protons in near-Earth space, based on all the particle data sets available at that time. Since the first version, the models were updated to AE-8 and AP-8 in the 1970s, documented by Sawyer and Vette [1976] and Vette [1991], of which both are available for solar minimum and solar maximum conditions based on the geomagnetic coordinates B/B_0 and L. These models remained unchanged for several decades, while the volume of data to further inform them increased a great deal, and consideration went into the updates required for AE-9 and AP-9, especially the changes described in the report by Ginet and O'Brien [2009], which have been replicated in table 1.2.

AE-8 and AP-8 were also relatively poor at describing the particle environment in the radiation belts, and in particular the SAA. Daly et al. [1996] has pointed out the difficulties associated with areas in the models where the gradient of the flux is high. Despite these shortcomings, the AE-8 and AP-8 models were the most-used for predicting magnetospheric radiation environments for several decades.

In 2013 the AE-9 and AP-9 were released to the public as described in Ginet et al. [2013] with the aim of improving some of the weaker areas of AE-8 and AP-8

Table 1.2: Prioritised requirements for an updated successor of the models AE-8 and AP-8 from Ginet and O'Brien [2009]. The orbital environments that these effects are relevant to are low Earth orbit (LEO), medium Earth orbit (MEO), geostationary orbit (GEO) and highly elliptical orbit (HEO).

Priority	**Population**	**Energy**	**Location**
1	Protons	>10 MeV (>80 MeV)	LEO & MEO
2	Electrons	>1 MeV	LEO, MEO & GEO
3	Plasma	30 keV to 100 keV (30 eV to 5 keV)	LEO, MEO, HEO & GEO
4	Electrons	100 keV to 1 MeV	MEO & GEO
5	Protons	0.1 MeV to 10 MeV (5 MeV to 10 MeV)	LEO, MEO & GEO

described in table 1.2. The new models include uncertainties using Monte-Carlo modelling which allows the user to weigh risks against the system design. It also includes many of the new datasets which has made it possible to increase the spatial resolution of the models.

1.4 Other planetary environments

Several other planets in the solar system also have a magnetosphere, in particular Mercury, Saturn and Jupiter, although the strength and spatial scale of these magnetospheres varies by orders of magnitude between them - Jupiter's is the largest structure in the solar system. Figure 1.14 illustrates this difference in extent of these magnetospheres, caused by both the strength of the planetary magnetic field and the influence of the solar wind at these distances from the Sun.

Each planet's magnetosphere plays a large part in the influence of these events on the atmosphere.

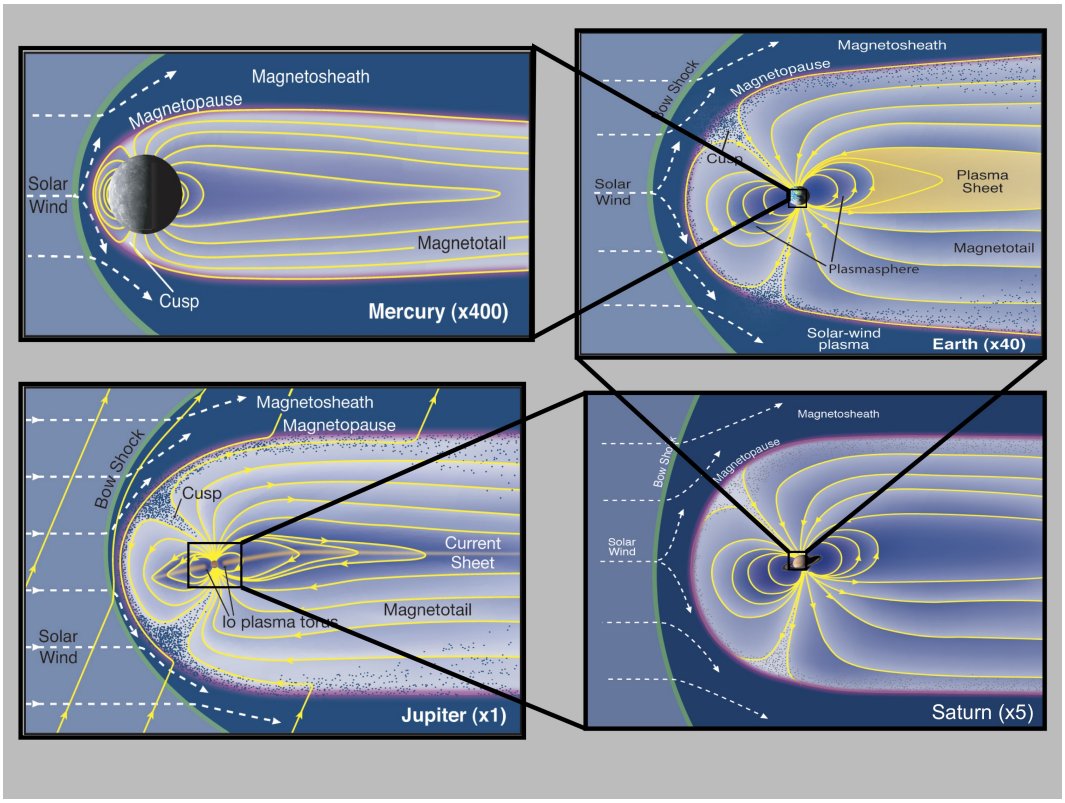


Figure 1.14: Comparison of planetary magnetospheres, from Fran Bagenal and Steve Bartlett at http://lasp.colorado.edu/home/mop/

The space environment near Mercury is characterised by a small planetary magnetosphere and almost no atmosphere; stable radiation belts do not exist. The few probes that have investigated the plasma environment at Mercury have concluded that the magnetosphere, although small, is still larger than had been expected, and exhibits similar behaviour to Earth's magnetosphere, but on a smaller scale in size and time. This and other planetary magnetospheres were described in the review by Russell [1993].

European Space Agency (ESA)'s next planned Hermean mission is Bepi-Colombo, due to launch in 2018; its scientific objectives include the investigation of the planetary exosphere and magnetosphere, described by Benkhoff et al. [2010]. While the scientific objectives of Bepi-Colombo and Mercury Surface, Space Environment Geochemistry and Ranging (MESSENGER) are very similar, the instrument suite on Bepi-Colombo is designed to gather a larger range of data, through three separate spacecraft - a lander and two orbiters at significantly different altitudes. The Mercury Plasma Particle Experiment (MPPE) on the magnetospheric orbiter part of the mission contains the High Energy Particle instrument (HEP) instrument, which is designed to detect electrons between tens and hundreds of keV, although the planetary orbiter, in a lower-altitude more circular orbit, will carry no energetic particle instruments.

The magnetospheres of Saturn and Jupiter are many times larger than the Earth's, and have clear radiation belts and visible auroras; auroras have also been observed on Uranus and Neptune. Jupiter's radiation environment in particular is highly harsh and dynamic. Three of the Galilean moons - Io, Europa and Ganymede - interact strongly with the planetary magnetosphere, and Io's volcanic emissions add a torus of charged material around Jupiter. Indeed, Io has been observed to have auroras of its own. The Jovian system has particularly high radiation levels because of the very strong magnetic field and spacecraft in its vicinity must be particularly well-protected against radiation damage.

Planetary missions such as the Cassini-Huygens mission to Saturn have previ-

ously been developed to include measurements of particles and magnetic fields at and around each of these planets. Multiple missions have performed flybys of Jupiter, often as a slingshot in the mission, including the Pioneer and Voyager probes, Ulysses and New Horizons. Others have reached Jupiter as their final aim, including the Galileo mission, which was launched in 1989. This carried, among other instruments, the Energetic Particles Detector (EPD) - a telescope design including a time-of-flight component.

Data on the Jovian environment has been collected from the Cassini mission flyby in 2000 and the New Horizons flyby in 2007. Since these, Juno was launched in 2011, carrying, among other instruments, the Juno Energetic Particle Detector Instrument (JEDI) to "explore and characterise the three-dimensional magnetosphere and auroras", arriving at and and beginning to orbit Jupiter in July 2016. Juno's first results include imaging the southern aurora at three wavelengths.

The JUpiter ICy moons Explorer (JUICE) mission to Jupiter has been selected by ESA as the next large-scale mission to explore the Jovian system and investigate the environments of Europa, Ganymede and Callisto. The aim is to launch the mission in early 2022, in which case it will arrive at the Jovian system in January 2030, described in Dougherty [2011]. Its instrumentation suite will include several experiments for in-situ and remote sensing of the plasma environment. Future data hope to help further characterise and explain the atmospheric and magnetospheric behaviour of the planet and moon system.

1.5 Platforms for space plasma research

The lower thermosphere has had relatively limited direct measurements taken of its properties, since it is comparatively inaccessible for long-term in-situ measurements both to orbiting and ground-based systems. However, there have been results from ground-based instrumentation, such as ionosondes, from sounding rockets and from spacecraft travelling in elliptical orbits with low-altitude perigees. These have given

a good indication of the conditions in the upper atmosphere; current and future missions aim to further this information with multi-position measurements over both large and small scales.

Small satellites in particular are likely to become used in more remote and challenging environments in the future, for many reasons: lower cost missions, faster deployment times, higher acceptable risk allowing the use of more innovative technologies as well as new scientific possibilities from multi-point measurements. The technology required for advanced operation, such as propulsion and deep space communication are developing fast and will become available in future years. At the same time scientific sensors are miniaturised and will become more suitable for integration into small satellites.

Some examples of small satellites that are currently proposed or planned for upcoming and future missions are given in section 1.5.3 and section 1.5.4.

1.5.1 Ground-based instrumentation

Ground-based instrumentation is preferred where possible; in addition to the greater flexibility in resource budgets, the hardware is far more accessible for repair, recalibration and overall traceability. Plasma science has benefited from both remote sensing and in-situ ground-based instrumentation throughout its history, from the very earliest auroral observations with the naked eye and solar disk observations with the aid of camera obscuras. Current plasma research using ground-based remote sensing technologies is well-developed, using a range of both active and passive methods distributed over multiple geographical locations.

Some of these methods, for example the fluxgate magnetometers described below, are also widely used for space-based plasma research, alongside particle measurements.

Ionospheric radar

Radar pulses can be used to take measurements over a cross-section of the atmosphere and ionosphere, aiming in particular to measure disturbances attributable to solar influences. Direct measurements include the electron density, electron and ion temperature, as well as ion composition and plasma velocity and the ion-neutral collision frequency, while the data can also be used to derive other temperatures and densities.

Incoherent scatter radar works by measuring the scattering of high-frequency radio waves in the acoustic waves within charged particles within ionospheric plasma. For example, the European Incoherent Scatter (EISCAT) radar which operates four radar sites in northern Scandinavia, described by Rishbeth and Williams [1985], has been operational since 1981 for some radar frequencies and since 1985 for the full range of frequencies. The very low power of the scattered signal requires very high transmitting powers of megawatts and large antennas to both transmit and receive signals.

A similar technique is used in ground-based ionosondes: a low frequency radar signal of a few MHz is used to identify layers in the ionosphere by measuring the return signal from an antenna pointing towards zenith. Cross-calibration between incoherent scatter radars and ionosondes is used to improve the accuracy of both systems.

Magnetometers

Ground-based magnetometers are often used to add additional information to space-borne instruments, for example for the Time History of Events and Macroscale Interactions during Substorms (THEMIS) mission Russell et al. [2008]. Chains of ground-based, low-cost flux-gate magnetometers have been constructed to be able to correlate the data from the five THEMIS spacecraft to ground based measurements during events. These magnetometers are sampling at 2 Hz and are synchronised using Global Positioning System (GPS) while uploading the scientific data at near-

realtime using the internet. A resolution of 10 pT is achieved, compared to 3 pT for the satellite instruments which were described by Auster et al. [2008].

Other observations

Other, essentially passive, measurements are used to add to these data. Ground based aurora observation is used to monitor magnetic substorms, for example the ground-based all-sky-imager array in the THEMIS mission described by Mende et al. [2008], which means the full pattern of visible auroras across the auroral regions of North America can be pieced together.

In addition, the total electron content in the ionosphere can be estimated using carrier phase delays of radio signals, for example of the GPS, by Mannucci et al. [1998] and others. By using the existing GPS ground station network it is possible to identify ionospheric features of about 500 km size.

1.5.2 Missions in near-Earth space

When ground-based instruments have reached the limit of what can be measured of the space environment from the Earth, there remains a large amount of data that may only be collected *in-situ*.

On these space missions, the resource budgets such as size, mass, power and telemetry may be highly restricted due to the extreme nature of the environment or the costs involved in launching larger and more complex spacecraft.

The low Earth orbit environment holds its own set of challenges for spacecraft. Missions here exchange relative safety from space weather events for the difficulties associated with atmospheric drag, and lower power requirements for telemetry for shorter pass times over ground stations.

Since the relatively high density of the thermosphere will significantly reduce the lifetime of the mission, most in-situ investigations aimed at the lowest orbital altitudes have fallen into one of several categories, described below.

Sounding rockets

Sub-orbital rockets are often used for microgravity research payloads, but have also been used for many atmospheric and ionospheric instruments, with apogees between altitudes of around 100 km to 1500 km. Such campaigns are limited to a maximum time period of a few tens of minutes in which to take measurements and are somewhat limited spatially by where they can be launched from or land.

Rockets have been very useful in observing small spatial- and time-scale changes in auroral events, in very many studies over multiple decades.

Elliptical orbits

A spacecraft with a highly elliptical orbit will only spend a small portion of its orbit at low altitudes. While such a spacecraft will eventually be de-orbited because of drag, the time period over which measurements can be taken is far longer than for a circular orbit. A mission with this type of orbit has the capability of repeating similar measurements over a reasonably long timeframe, with some loss in time resolution. For example, the Atmospheric Explorer series of satellites each had perigees of around 150 km to 250 km with a range of inclinations, and made some of the only atmospheric measurements at such low altitudes.

Boosted spacecraft

For low-orbiting satellites which require longer lifetimes, for example the ISS, it is a commonly-employed practice to allow the altitude of the orbit to decay slightly over time and boost it to a higher orbit periodically. Alternatively, the Gravity field and steady-state Ocean Circulation Explorer (GOCE), which orbited at a very low altitude of 250 km, used an ion thruster to continually act against the slight air resistance and produce the exceptionally smooth trajectory required by its sensitive gravitational field instruments.

CubeSats and nanosats

The CubeSat standard was developed by Robert Twiggs at Stanford University and Jordi Puig-Suari at California Polytechnic in 1999, and the standard's latest documentation is provided by Munakata [2009]. The form factor is a 10 cm-side cube, with a mass of no more than 1.33 kg. Since the constraint of such missions is often cost and launch availability, the purpose of such standardisation is to assist in reducing the cost of the launch and deployment and also reducing the risk to both the CubeSat and the main mission.

The first CubeSats were used primarily for educational projects, but the next generation was soon used widely for technology demonstrations. In the past decade their further uses for scientific and imaging payloads have come to the fore, with the first set of commercial companies offering CubeSat based services.

CubeSats are often used in relatively low-altitude orbits, since their simplified missions with low cost and low development time are typically suitable for shorter missions before the spacecraft are deorbited by atmospheric drag. The less expensive, less powerful launchers required to launch to low altitudes are in keeping with the smaller budgets of these missions.CubeSats also allow flying simplified and cheaper instrumentation with a shorter planned lifetime and a higher failure risk, accepting that these will not be able to withstand longer periods in a high radiation environment.

A strong advantage of CubeSats in LEO is the ease of launching multiple spacecraft at one time. A constellation mission refers to one in which the spacecraft are manoeuvred to provide the coverage needed - whether measurements close together or globally. A swarm mission does not use propulsion to control the relative positions of the measurements, but relies on the number of spacecraft to cover the measurements at a wide enough spread to produce a good coverage of the area of interest.

Swarm missions are a suitable and attractive proposal for a mission composed of many small low-cost CubeSats; here the loss of a fraction of the total number

is acceptable, and with enough units, propulsion is unnecessary to keep a good coverage of the planet.

When the measurements of interest will be collected over a relatively short timeframe, these would be the most efficient at collecting high-resolution data; multi-point measurements increase the spatial resolution and decrease the time between measurements in a single place.

An example of a swarm mission is QB50: a planned constellation of 50 CubeSats due to be launched in early 2017, with the aim of providing a network of platforms for miniature technologies which can be used for space weather research. The instrumentation plans for these satellites are described in the sensor selection working group report by Smith et al. [2012]. Around 40 of the units will be double CubeSats and will contain standardised systems and instrumentation, along with the possibility of individual institutions adding additional miniature experiments or demonstrations in the small volume remaining. All of the satellites will be launched to initial circular orbits at an altitude of 320 km - just below that of the ISS. Their orbits are expected to decay through atmospheric drag to 90 km altitude in the mission's lifetime of around 3 months before re-entry. The orbits may be at an inclination of up to 79°. This project is unique in using a network of orbiting satellites at such low altitudes; the low cost of CubeSats leads to the ability to invest in a mission with such a short lifetime.

1.5.3 Beyond the magnetosphere

Missions to beyond the Earth and its magnetic field's influence have previously been limited to large scale missions such as those mentioned briefly at the end of section 1.4. However, the increase in capabilities of CubeSats and other small satellites over the past decade has been such that preparations are now being made for the first lunar and planetary CubeSat missions. In the past, such small-scale scientific missions were judged to be too risky and unable to justify the cost of launch or transit but although the concept is in its infancy, some missions have

been conceived which will become some of the first to the Moon, to interplanetary space and beyond.

For example, the proposed Interplanetary NanoSpacecraft Pathfinder In Relevant Environment (INSPIRE) mission from Jet Propulsion Laboratory (JPL) described by Klesh et al. [2013] plans to launch two CubeSats beyond the magnetosphere as a demonstration of the possibility of system performance and communication in a more remote environment. This type of mission is flexible enough to be launched as a piggyback on any appropriate interplanetary mission launching in the next few years.

The possibility of using small satellites with a CubeSat form factor as a separate sensor or relay spacecraft in remote environments has been discussed. For example, the Interior Exploration using Seismic Investigations, Geodesy and Heat Transport (InSight) mission to Mars, which is planned to launch in 2018, plans to include the Mars Cube One (MarCO) CubeSat, and a second identical to it for redundancy, which will provide a real-time data relay during descent and landing of the lander.

CubeSats are also proposed for scientific research; the ESA is currently selecting 6 CubeSat units in a currently undefined configuration to be deployed from the Asteroid Impact Mission (AIM) spacecraft to study an asteroid, described by Carnelli et al.. Another current mission concept is JPL's design for a lunar CubeSat mission, Lunar Flashlight, to launch in 2018, described by Cohen et al. [2013]. Here a 6U CubeSat will be the first CubeSat to reach the Moon, landing at the lunar pole, where it will illuminate and analyse the surface, to estimate size and composition of ice deposits. It is also hoped that such a mission could be used as a precursor to future larger-scale missions, since pre-determining the properties of the local surface could provide information to the plans for a future lander.

NASAs Exploration Mission 1 (EM-1) mission will host 13 CubeSats mission on the first flight of the Space Launch System (SLS) to the Moon in 2018.

1.5.4 Small satellite flight opportunities

CubeSats and other small satellites are gaining momentum for scientific research, with applications in many fields including Earth observation and space plasma research. They provide a strong opportunity as a platform, at first to demonstrate the possibilities of a miniaturised instrument, but they are also capable of multiple mission configurations which can allow scientific measurements in a range of environments, positions and timescales.

Several flight opportunities for small satellites will come up in the next few years, both in LEO and beyond. Calls for bids for flight on launch opportunities, such as those described in section 1.5.3, are relatively frequent, especially for educational CubeSats. In addition to this, opportunities and companies exist for the purpose of launching small satellites that meet the CubeSat standard. For example, NanoRacks provides CubeSat deployments from the ISS and Innovative Solutions In Space (ISIS) provides organisation and support for piggyback launches of CubeSats and other small to medium satellites.

Miniaturisation of plasma instrumentation is ongoing; multiple CubeSat-capable instruments both for fields and for particles have been developed, some examples of which are described in section 2.2.

The instrument described in this thesis adds additional capability to the family of miniaturised plasma instruments which could be used on CubeSats.

Chapter 2

Charged particle detection systems

The electron was discovered in 1897 and later the same year the mass to charge ratio was measured by Thomson [1897], but even before this, cathode rays of electrons in vacuum tubes had been observed as early as 1869 and named by Goldstein [1880]. The identification of further charged particles led to the development of more and more ways of detecting them.

At its simplest, a detector is affected by radiation and produces a useful electrical signal in response. In addition, it may have the capability to identify the directions, energies or species of incoming particles, either individually or statistics of the distribution.

2.1 The development of in-situ charged particle detectors for space plasmas

The history of charged particle detection pre-dates the space age by over 50 years, and is an area of instrumentation that has a hugely varied range of uses across many

fields of science and technology. From the earliest detectors of particle emission from radioactive elements developed over a century ago there have been many further developments, but even the earliest and simplest of concepts are still in use today - in space, but also in high-energy physics, medicine, nuclear power, materials science and other industries.

Early research into the properties of materials in response to particle radiation was advancing rapidly by the turn of the twentieth century for example through the work of Becquerel [1900] who investigated the range of alpha particles in solids, and of Rutherford [1899], who published his results on the stopping power of different materials and the absorption of different gases of uranium radiation.

Here is given an overview of *in-situ* detectors used for charged particles in space, both in plasma and at higher energies, and some of the related systems used to support the analysis of these particles.

2.1.1 Geiger-Müller tubes

The Geiger counter was developed in 1908 by Hans Geiger, along with Ernest Rutherford, but was only capable of detecting alpha particles. The Geiger-Müller (GM) tube was introduced by Geiger and his graduate student Walther Müller and published as Geiger and Müller [1928] and later papers.

The counter uses a high voltage across a gaseous medium to detect a single particle of ionising radiation. Once one gas particle is ionised, it triggers an avalanche of ionised particles which creates a signal that can be read out as an event.

GM tubes have been used as particle detectors on many early missions, such as Sputnik 2, the second spacecraft to orbit the Earth, and Explorer 1, the first US satellite. Data from these instruments allowed Van Allen [1958] and Van Allen [1959] to detect and identify the radiation belts for the first time.

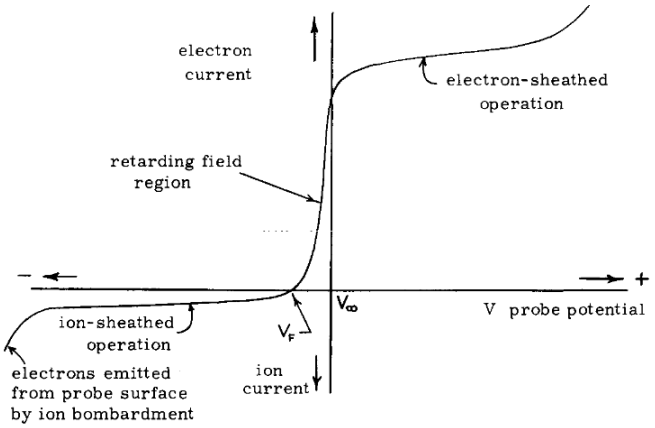
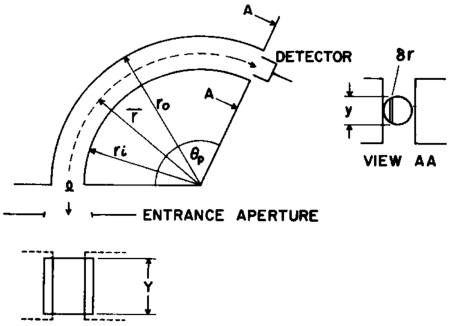


Figure 2.1: Langmuir probe characteristic from De Leeuw [1963]. Ion current exaggerated.

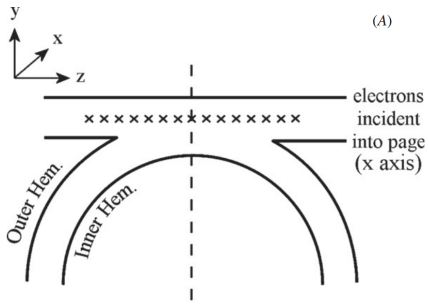
2.1.2 Langmuir probes

A Langmuir probe is, at its simplest, one or more exposed metal electrodes inserted into a plasma away from the spacecraft which, by varying their electric potential, can use current measurements to derive plasma properties. The measured I-V characteristic of the Debye sheath around the probe can be used to derive particle population characteristics, such as the ion and electron density and the electron temperature in thermal plasma, based on the knowledge that electrons are moving far faster than protons and ions with a basic response curve as shown in figure 2.1. Early theoretical and experimental development of Langmuir probes included that of Mott-Smith and Langmuir [1926].

While Langmuir probes on large-scale missions can be positioned on booms several metres long, a miniaturised Langmuir probe will fly on several of the CubeSats in the QB50 constellation.



(a) A cylindrical geometry from Johnstone [1972]



(b) A cross section of an hemispherical top hat geometry from Collinson and Kataria [2010]

Figure 2.2: Diagrams of (a) a cylindrical geometry and (b) a top hat geometry

2.1.3 Electrostatic analysers

Low energy electrons and ions can be purposely deflected by electric and magnetic fields within an instrument; this can be used to analyse the energy spectrum of charged particles, since the radius of the path in a constant field will be defined by the energy, mass and charge of the particle. In one dimension, a cylindrical geometry is used as shown in figure 2.2 (a), but a common design with a much larger geometric factor is a spherical shape known as a top hat analyser, a cross section of which is shown in figure 2.2(b). Top hat analysers are particularly powerful since they can be capable of measurements over nearly a full 4π range of angles, and energy and angle of particles can be swept through by varying the voltage at the entry and across the hemispherical surface.

Such analysers have been used in a large number of space missions, although they are restricted for smaller missions by large size and mass or low geometric factor, and would require high electric fields and, correspondingly, voltages to be usable for electrons above around 10 keV to 100 keV. However, some miniaturised examples are described in section 2.2.2.

2.1.4 Channel electron multipliers (CEMs) and MCPs

Common methods of detecting low-energy charged particles include CEMs and MCPs. A CEM is a tube structure with a voltage between the ends. A charged particle hitting the surface can produce secondary particle emission which is also accelerated and may itself produce further secondary electron emission to produce an avalanche of electrons and thereby a current pulse at the other side. An MCP is essentially a plate made of many small parallel CEMs; the high voltage (usually in the range of kV) is applied across the plate and incoming radiation produces a detectable signal at the other side.

MCPs and CEMs have often been used in conjunction with electrostatic detectors; the choice between the two and of the type is made depending on the angular resolution and gain required.

2.1.5 Scintillation detectors and photomultipliers

Scintillators are made from materials that produce photons when impinged upon by an incident charged particle. The incident particle causes the scintillator to re-emit its absorbed energy in the form of light.

A scintillation detector is usually based on a photomultiplier tube (PMT) or photodiode to detect the light emitted from a scintillator. Both of these types of photon detector make use of the photoelectric effect to produce a current from incident light. A PMT contains a photocathode in a vacuum which releases electrons when illuminated by light. This electron signal is then accelerated and amplified by a series of dynodes at increasing potentials to produce an easily read-out output current. A photodiode is a semiconductor device which also converts light into current. Incident photons create electron-hole pairs which are swept apart by the electric field and therefore create a current. An avalanche photodiode is run with a high reverse bias voltage which causes carrier multiplication within the device and therefore has intrinsic gain. Both a PMT and and avalanche photodiode require

a high voltage supply. Each of these detectors produces a current signal to be processed and analysed by further electronics.

A part of the detector Energetic and Relativistic Nuclei and Electron instrument (ERNE) instrument on the SOlar and Heliospheric Observatory (SOHO) is based on a scintillator and photodiode combination and is described by Torsti et al. [1995].

2.1.6 Solid state detectors

A solid state detector uses a solid semiconductor as the detection medium. The displaced electrons within the lattice structure of the material are collected on electrodes and the signal is measured.

Since these detectors will be the focus of the research in this thesis, and the behaviour and interactions of silicon in particular are important, more description of the properties and workings of solid state detectors and their designs will be given in section 2.3.

2.2 Examples of missions including small-scale charged particle detector instruments

Each region of outer space within, and beyond, our solar system has its own characteristics and interest for plasma science. Spacecraft have travelled to the magnetospheres of Earth and of the other planets, and to different areas of the heliosphere, where they can observe the solar wind, its behaviour and also how it reacts to energetic solar events and eruptions. Voyager 1 and 2 have even reached the termination shock, where the interaction between the interstellar medium and the edge of the Sun's influence causes a shock in the heliospheric plasma, and are currently studying the heliosheath beyond. Almost all of these missions have involved some form of *in-situ* plasma detector or analyser, even from the earliest days of space travel. On the very earliest missions, instruments were by necessity simple and compact to

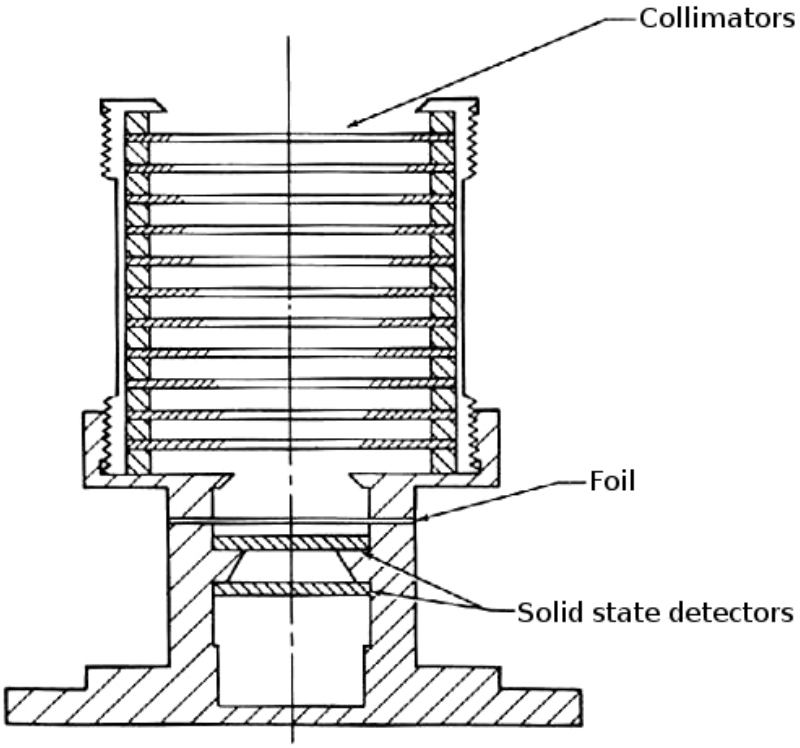


Figure 2.3: Layout of one of the two solid state telescopes on the Apollo 16 subsatellite from Anderson et al. [1974]

match the launch capabilities and fit in with other necessarily large systems. One early example included the Lunar Particles Shadows and Boundary Layer instruments flown on the Apollo 15 and 16 subsatellites in 1971 and 1972. They contained some of the first small-scale energetic particle telescopes, which had two solid state detectors; they are shown schematically in figure 2.3.

As launch and mission capabilities grew and the scope of space science increased, instruments and spacecraft needed to increase in complexity and be on large scales to produce useful data. For example, electrostatic analyser instrument packages, used for electron detection and magnetospheric research, have undergone significant

changes. The earliest electrostatic analysers on missions in the 1960s were relatively simple deflector plates. The top hat family of instruments has been used in missions from sounding rockets as early as the 1970s to demonstrate the concept. In the decades since, the design evolved in complexity to investigate electron populations in the magnetosphere in ever more sophisticated ways.

One of the earliest top hat instruments to fly in orbit was the package on the Ion Release Module (IRM) spacecraft which was part of the Active Magnetospheric Particle Tracer Explorers (AMPTE) mission launched in 1984. Two top hats approximately 10 cm in diameter measured electrons between 15 eV to 30 000 eV and ions between 20 eV/q to 40 000 eV/q from Paschmann et al. [1985]. Data processing and allocated telemetry budget limited the data products: the instrument was required to use variable measurement and compression schemes to maximise its scientific return. The Plasma Electron And Current Experiment (PEACE) instrument on Cluster, which successfully launched in 2000, was a development from this. As with AMPTE it used 2 top-hat analysers, but at a higher sampling rate. It also had a wider energy range down to 0.7 eV. PEACE was somewhat larger, weighing 5.4 kg and using an average of 4.7 W of power. The Magnetospheric Multiscale (MMS) mission launched in 2015 with the aim of sampling electron and ion populations every 150 ms and SI30ms respectively. It carried 8 top hat analysers per spacecraft, which allowed continuous 360° measurements even at a slow spacecraft spin rate. Because of this repetition, the instrument set weighed 54.5 kg and, since each top hat required multiple high voltage supplies as well as support and processing circuitry, consumed between 50 W to 67 W depending on the operation mode.

However, this trend has triggered an additional shift in more recent years: the emergence of the use of nanosatellites as scientific probes. Technological improvements have opened possibilities for highly-capable smaller instruments intended for flight on a wide range of mission scales, and CubeSats and other small satellites are evolving into platforms for a range of scientific instruments. Figure 2.4 shows developments of the capabilities of CubeSats by the number of launches per year for

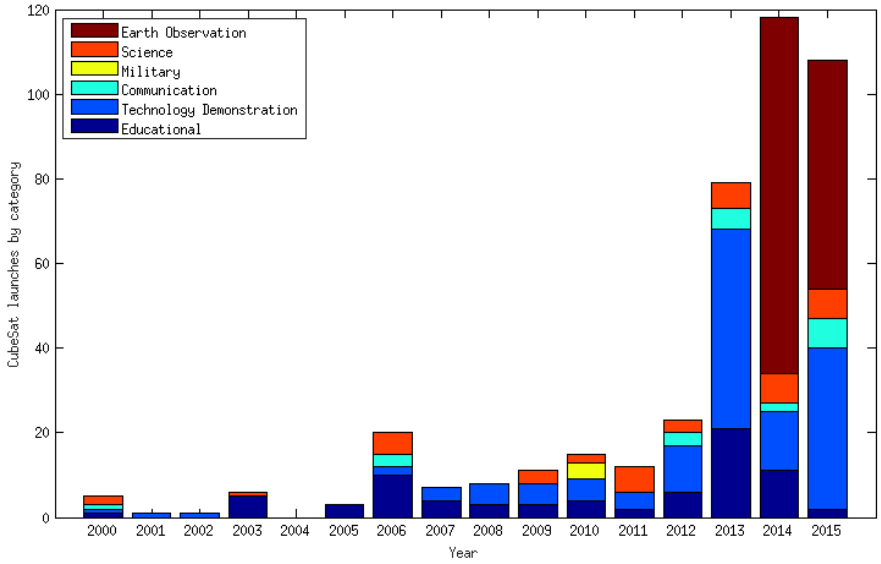


Figure 2.4: Number of CubeSat launches each year since their development by category of mission. Data from M. Swartwout.

each field in the the past two decades up to 2015.

The trend towards small satellites and CubeSats provides the possibility of many smaller platforms for new instruments, while on the other hand significantly reducing the overall available resources that can be used by an instrument. Attempts have been made to improve the state of the art capabilities of small-scale plasma and energetic particle instruments, or to produce instruments capable of compromise between scientific requirements and minimal resource usage.

The examples given in this section illustrate the range of attempts to miniaturise existing instrument design concepts to meet these emerging mission requirements and take advantages of the new possibilities for space science they offer.

2.2.1 Miniaturised solid state detectors

A solid detection medium is a major advantage for high-energy electron detectors where a priority is miniaturisation, because the density of the medium is around three orders of magnitude higher than in a gas. Since the detection of ionising radiation is based on interactions within the detector, the shorter stopping distance and path length in between interactions in a solid means a far smaller detector can be used to detect or absorb the particles of interest.

Additionally, solid state detectors usually have relatively low voltage requirements, have a wide energy range and may be capable of distinguishing types of particle. They can be used to produce versatile instruments which have the potential to reduce the number of instruments needed on a spacecraft. They are also a well established technology - small versions can usually be purchased as Commercial off-the-shelf (COTS) components, which reduces the cost and instrument development time.

However, certain solid state detectors may require cooling systems and for position sensitive detectors readout electronics may be more complex, especially for the retrieval of detailed track information or particle type. They are also susceptible to radiation damage, but it is often possible to restore some of the performance through annealing.

The lower energy limits of solid state detectors are determined by the thickness and structure of the dead layer. For very thin entrance windows, the detector becomes increasingly sensitive to light which makes particle identification difficult in all but the very darkest environments. Typically, these sensors would be used for higher particle energy ranges than MCPs since some dead layer which lower-energy particles cannot penetrate is unavoidable.

Standard Radiation Environment Monitor (SREM)

SREM is ESA's Standard Radiation Environment Monitor, designed with the aim of producing a standardised low-resource radiation monitor with a mass of 2.6 kg

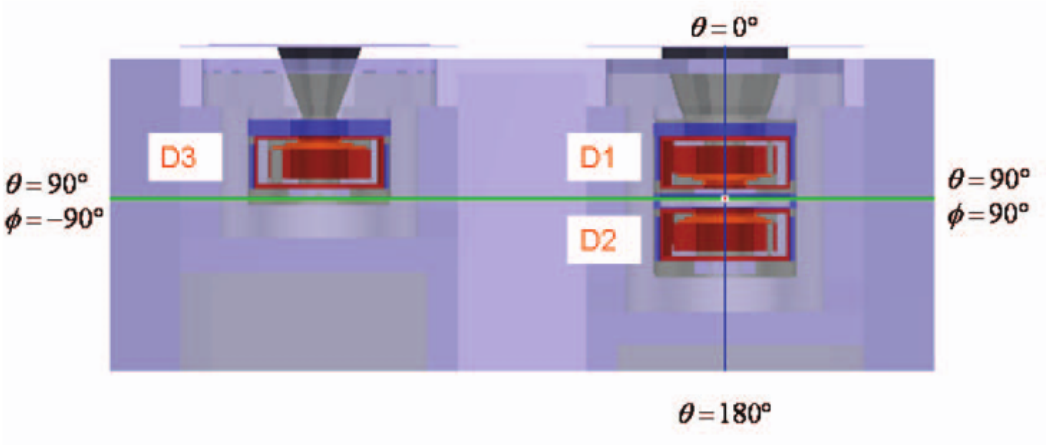


Figure 2.5: The SREM setup from Siegl [2009]

which could be added as an additional sensor to multiple missions as a standalone unit. The design and calibration was described in the master thesis by Siegl [2009]; three solid state detectors are arranged in a telescope pair and a single detector setup, as shown in figure 2.5.

So far, SREM, or a slightly modified version, has been flown on at least ten orbiting and interplanetary spacecraft.

Micro Radiation Environment Monitor (MuREM) and Radiation Monitor (RM)

Both MuREM and Surrey Satellite Technology Limited (SSTL)'s RM (Taylor et al. [2011]) have flown on TechDemoSat-1 (TDS-1) as a demonstration of miniaturised instruments capable of providing information to a spacecraft system both to allow protection in the case of dangerous radiation environments and to provide scientific data on space environments. Both devices carry radiation field effect transistor (RADFET) dosimeters, photodiodes and diodes used as particle detectors to measure the environment's likely effect on on-board electronics.

SSTL RM is also connected to further RADFET sensors on the spacecraft, whilst

MuREM carries a radiation effects payload of devices for monitoring damage from energetic proton or ion events.

MuREM is prepared for the use in CubeSats and uses three PC/104 boards, while SSTL RM fits within a 6 cm × 6 cm × 10 cm envelope.

Relativistic Electron Proton Telescope integrated little experiment (REPTile)

The REPTile (Schiller et al. [2010]) is a scaled-down version of a larger experiment, Relativistic Electron Proton Telescope (REPT), which was launched on the dual Van Allen Probes in August 2012. For each instrument, the design consists of a stack of detectors, a beryllium window and a collimator with a series of baffles, but the resources are significantly reduced: REPT's 13.4 kg was reduced to just over 1 kg in the case of REPTile. The setup is optimised for high-energy particles in the radiation belt, with energies of 0.5 MeV to 3 MeV for electrons and 10 MeV to 40 MeV for protons. It was launched in a 3U CubeSat format into a circular orbit at around 600 km altitude on the Colorado Student Space Weather Experiment (CSSWE) CubeSat mission in 2012.

The design which is shown in figure 2.6 is 6.05 cm in length and 6 cm in diameter, and uses <1 W of power, which are likely to be approximate limits for a CubeSat instrument.

The Highly Miniaturised Radiation Monitor

The Highly Miniaturised Radiation Monitor (HMRM) is a highly-integrated telescope-based instrument for dosimetry and identification of energetic charged particles in a very small package. Figure 2.7 from Mitchell et al. [2014] shows the alignment of the active pixel sensor (APS) telescope surrounded by application-specific integrated circuit (ASIC)-based electronics; the instrument inside its housing is 1.7 mm × 2.4 mm × 2.2 mm. Such miniaturisation is only possible with highly-integrated circuitry and significant development time.

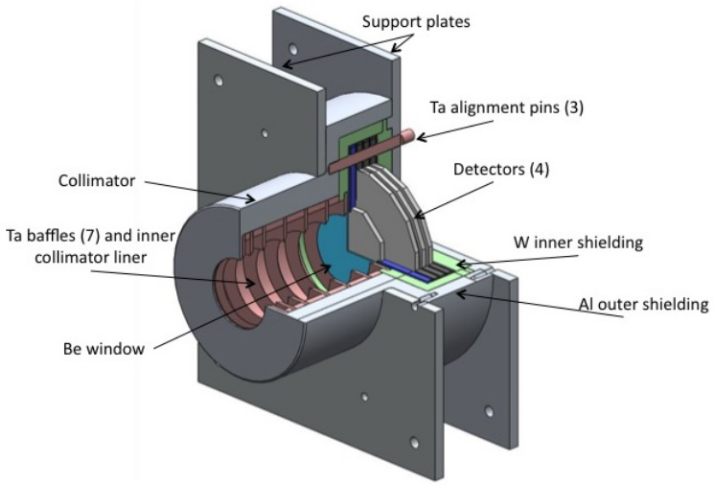


Figure 2.6: The REPTile instrument from Schiller et al. [2010]

A technology demonstration of HMRM was flown on TDS-1, and it is hoped that future versions may be flown as secondary payloads on future missions, especially those which would benefit from real-time dosimetry.

2.2.2 Miniaturised electrostatic detectors

The potential of small or simple detectors can be enhanced by the use of particle optics. Deflecting electric fields within an instrument need to be stronger on smaller scales and with higher particle energies, so the production of of high voltages must be traded off against the scale of the instrument.

Conceptual And Tiny Spectrometer (CATS)

The CATS instrument was designed and prototyped at MSSL by Bedington et al. [2012], as a highly miniaturised cylindrical analyser, with concentric channels which had alternating electric field directions and therefore was capable of distinguishing

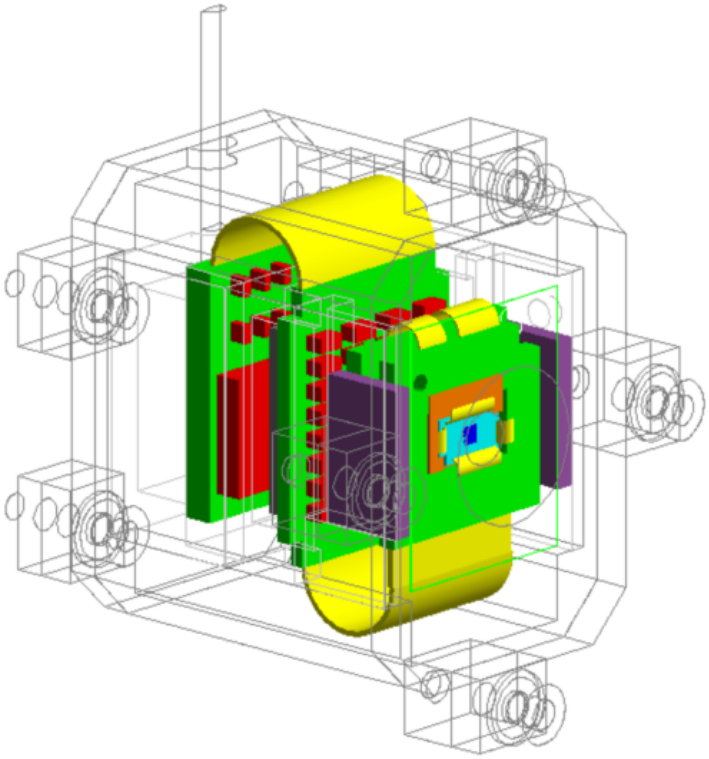


Figure 2.7: The Geant4 geometry model of the HMRM instrument from Mitchell et al. [2014] showing detectors, ASICs and flexible wiring

between both electrons and positive ions at five energy bands each. This instrument was tested and calibrated with a CEM and a back-illuminated CCD in the laboratory and could be optimised for a space environment by selecting a different detector.

Other attempts have been made in curved plate analyser research to advance miniaturised top hat detectors such as Miniature Optimized Smart Sensor (MOSS) and Advanced Miniature Plasma Spectrometer (AMPS), which are described in Funsten and McComas [1998] although these have not flown.

SupraThermal Electrons, Ions and Neutrals (STEIN)

Cubesat for Ion, Neutral, Electron, Magnetic fields (CINEMA) contains STEIN developed by SSL (Glaser et al. [2009]), along with a magnetometer provided by Imperial College London. The satellite was launched on 13 September 2012 as part of NASA's ELaNa programme.

STEIN is designed to measure particles in the 2 keV to 300 keV range and differentiate between different types of particles that have energies of up to 20 keV. Using the same four-pixel solid state detector as that used in the STE instrument on STEREO, an electric field and an aperture, high energy particles reach the central pixels and lower energy charged particles are deflected to the edges, as shown in figure 2.8.

2.2.3 Summary of miniaturised instruments

Table 2.1 shows a number of miniaturised detectors with their masses and energy detection ranges. In recent years, advances in electronic and detector technology, along with the uptake of the CubeSat standard, have made small instruments more and more feasible. Small low-resource instruments with high capabilities are now required to increase the potential of space research.

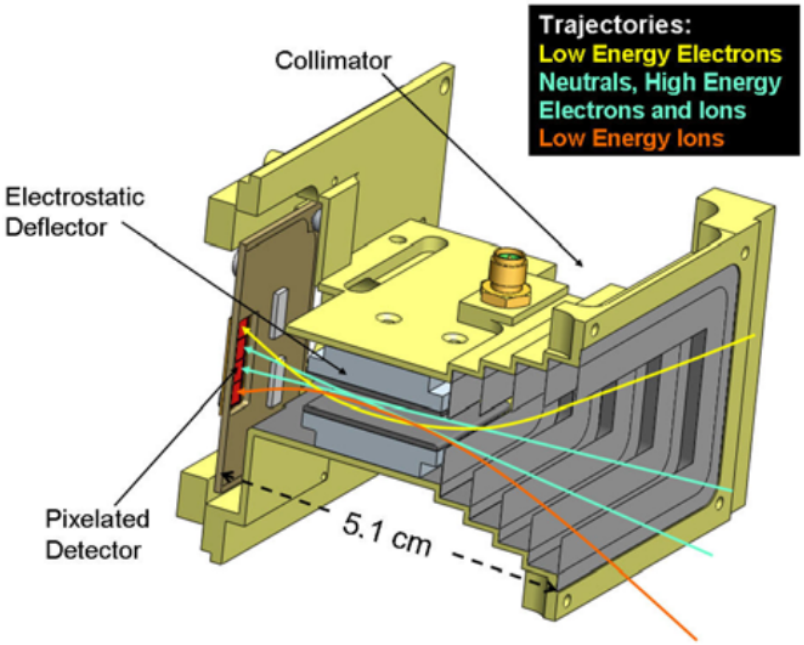


Figure 2.8: Description of the STEIN solid state detector on CINEMA from Glaser et al. [2009]

Table 2.1: Energetic particle instruments on small satellites

Instrument	Active years	Mission	Mass	Energy range	Citation
SREM	2000–	Various	2.6 kg	0.3 MeV to 15 MeV e^- 8 MeV to 800 MeV p	Siegl [2009] Mohammadzadeh et al. [2003]
REPTile	2012–2015	CSSWE	1.25 kg	500 keV to 3 MeV e^- 10 MeV to 40 MeV p	Schiller et al. [2010]
STEIN		CINEMA	0.4 kg	2 keV to 300 keV	Glaser et al. [2009]
MuREM	2013–	TDS-1	0.5 kg	All (dose)	Taylor et al. [2011]
SSTL RM	2013–	TDS-1	0.4 kg	All (dose)	Taylor et al. [2011]
HMRM	2013–	TDS-1	52 g	80 keV to 5.76 MeV e^- 1.41 MeV to 460 MeV p	Mitchell et al. [2014]

2.3 Solid state detectors

A solid state detector is, at its simplest, a piece of semiconductor capable of converting energy deposited in it by ionising radiation into an electrical signal which can be read out and analysed. Since the research described later in this thesis is focused on silicon detectors, a description of semiconductor operation processes is given in this section.

The first generation of semiconductor detectors, in the 1960s, was focused on simple monolithic detectors capable of energy spectroscopy; it was not until the 80s that position-sensitive devices became the focus of research.

In general a silicon diode detector refers to one using a direct form of signal production, as opposed to other semiconductor-based detectors, for example those based on scintillators, which detect secondary generated photons.

2.3.1 Semiconductor detector properties

In a crystalline semiconductor, electrons must exist within defined bands of energy states. Valence band electrons are bound to lattice sites around nuclei; conduction electrons are able to move through the crystal, and are able to conduct current through the material. In an insulator, the gap between these energy bands is large (generally >5 eV), while in a semiconductor the bandgap is around 1 eV.

When ionising radiation is incident on a silicon detector, electrons are knocked out of their position in the lattice, forming pairs of electron and hole charge carriers. The energy that must be deposited to form one pair varies by type of semiconductor and is usually taken to be

$$\epsilon = 2.8E_G + r\,(\hbar\omega) \tag{2.1}$$

where the first term covers both the intrinsic bandgap (E_G) and the residual kinetic energy ($1.8E_G$) and the second term represents optical phonon losses. r is an adjustable parameter which is a measure of the ratio between mean-free-paths of

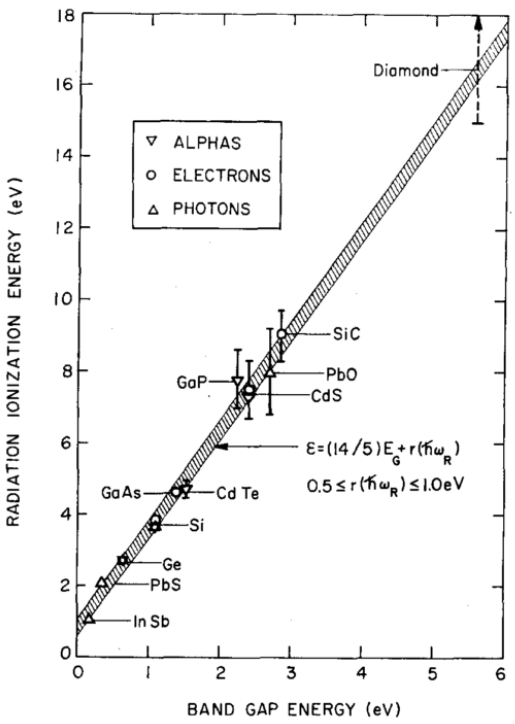


Figure 2.9: The relationship between ionisation and bandgap energies from Klein [1968]

ionising collisions and phonon emission, observed to be at a value of $r\,(\hbar\omega) \approx 0.6\ eV$ by, for example, Owens and Peacock [2004]. Other largely equivalent estimates have been made, but this, given by Klein [1968], is well-matched to the data and has been in use since the 1960s.

In the case of silicon this ionisation energy for an electron-hole pair is usually around 3.64 eV for most detector setups. Other detector materials are shown in the graph in figure 2.9 with the line from equation 2.1.

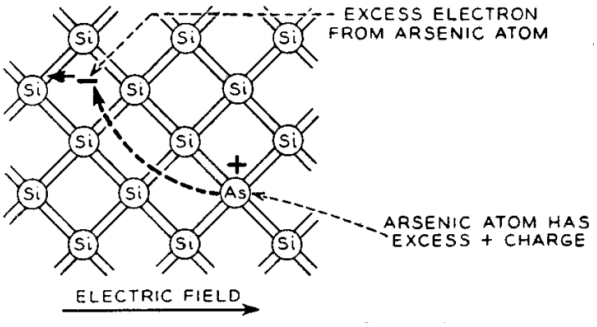


Figure 2.10: n-type silicon doped with an arsenic atom, detail from Shockley [1950]

Effect of dopants

Silicon detectors must be made of very high-purity silicon with a near-perfect crystalline structure. This is usually achieved by suspending a starter crystal in a bath of molten silicon and gradually, over a period of days, lifting it from the silicon to allow new layers of crystal structure to slowly form on the bottom of the column at the surface of the liquid. This technique was first documented by Czochralski [1918] as a new method for measuring the the crystallisation rate of metals. Teal and Little [1950] used the method to grow large single crystals of germanium.

Dopants can be added to change the structure and properties of a semiconductor. For example, replacing one of the silicon atoms in the lattice with a donor atom which has an additional valence electron (for example, lead, phosphorus, arsenic or tin) will leave one electron lightly bound to the impurity atom, since it cannot be accommodated in the valence band. This situation with a silicon lattice doped with a single arsenic atom is shown in figure 2.10. Since the majority charge carriers in this case are the negative electrons, this type of doped silicon is known as n-type and the dopant is called a donor.

In a similar way, a silicon atom replaced by a group 3 acceptor atom, such as boron, aluminium, gallium or indium, will remove a valance electron from the silicon leaving a hole, which produces p-type silicon. This situation, featuring a

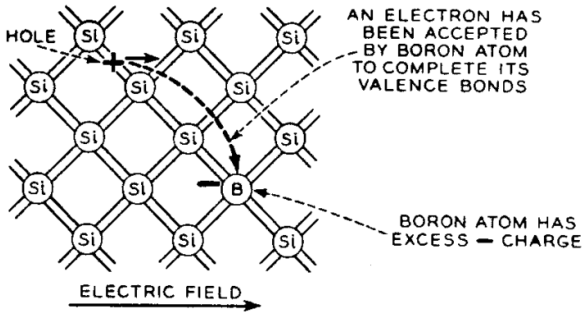


Figure 2.11: p-type silicon doped with a boron atom, detail from Shockley [1950]

boron dopant atom is shown in figure 2.11.

The effect on the lattice wavefunction from the donor or acceptor impurity spreads over many lattice sites and so by adding well-spaced impurities, a region of excess charge carriers is formed.

The method by which n- and p-doped layers are manufactured can be chosen to optimise the properties of a detector for an application. In the early years of device manufacturing crystals involving p-n junctions were grown using a modified Czochralski method described by Teal et al. [1951]. n- and p-doped regions can also be manufactured by allowing high-temperature dopant ions to diffuse into the surface layer, or for the highest level of accuracy and control, accelerating dopant ions to the necessary depth below the surface.

The depletion region

A silicon detector will contain both n- and p-doped material; the joining region between the two is called a pn-junction, and provides the properties necessary for radiation measurement.

The behaviour of the electric field by the pn-junction creates a depletion zone where mobile charge carriers have diffused away. The layer of the detector surrounding the join between the p- and n-doped layers contains very few excess electrons

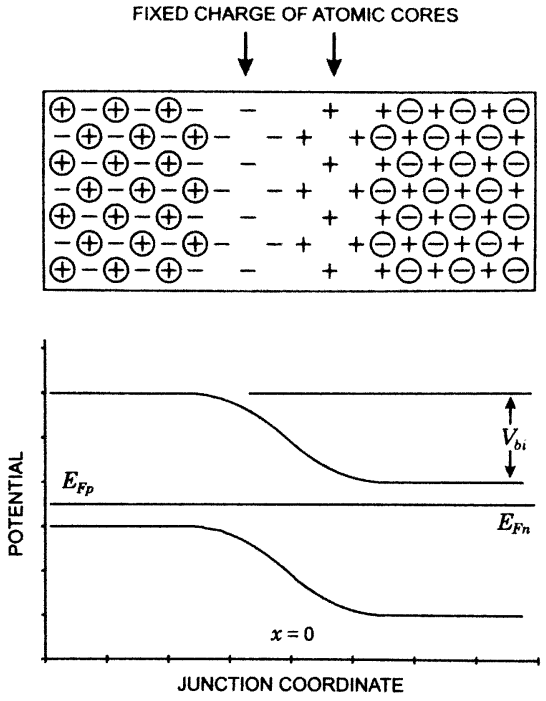


Figure 2.12: The structure of the depletion region and valence and conduction bands around a *pn*-junction from Spieler [2005]. E_F represents the Fermi level and V_{bi} is the potential difference created across the depletion region.

or holes. This diffusion creases a potential between the *p*- and *n*-regions, shown in figure 2.12.

In a detector, this is very useful. Incident ionising radiation with sufficient energy can produce an electron-hole pair in the lattice, indicating its presence. However, to take advantage of the properties of this depletion region and identify this electron-hole pair, a voltage must be applied. This voltage significantly alters the behaviour and composition of the depletion region.

Because of the higher mobility of electrons compared to holes, the *pn*-junction's current-voltage relationship is not symmetrical. The Shockley ideal diode equation

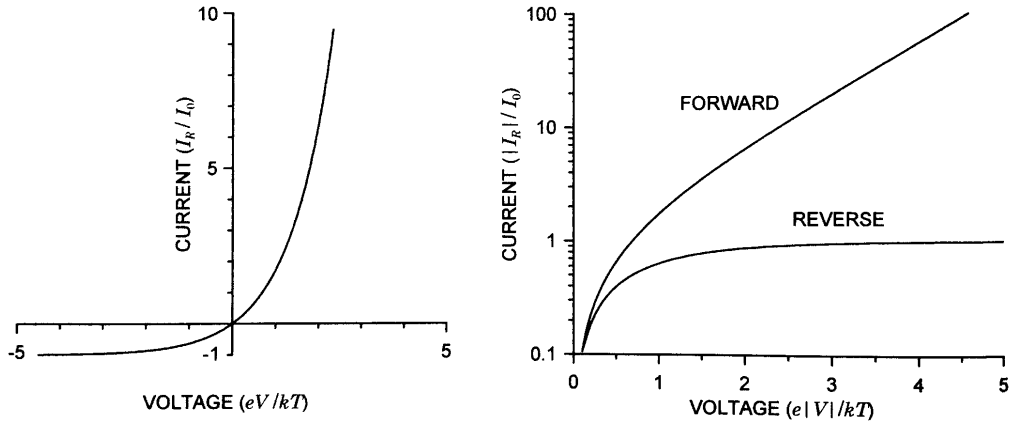


Figure 2.13: The current dependence on voltage of a semiconductor diode from Spieler [2005].

from Shockley [1949] is given by

$$I = I_0 \left(e^{\frac{eV}{kT}} - 1 \right). \tag{2.2}$$

This dependence is shown in figure 2.13. Under forward bias, the current increases rapidly, while under reverse bias the current quickly saturates. The saturation current depends on the material of the detector, specifically its bandgap, and defects or damage to its crystal structure.

Figure 2.14 shows how a bias voltage across a detector changes the movement of charge carriers and affects the potential of the p and n sides of the junction. Forward bias - where a positive voltage is applied to the p-doped material - lowers the potential difference and increases the movement of charge carriers across the junction. Reverse bias - where a positive voltage is applied to the n-doped material - raises the potential barrier and reduces the flow of charge carriers across the junction, which increases the width of the depletion region.

Reverse bias diodes form a region without mobile charge carriers. This creates

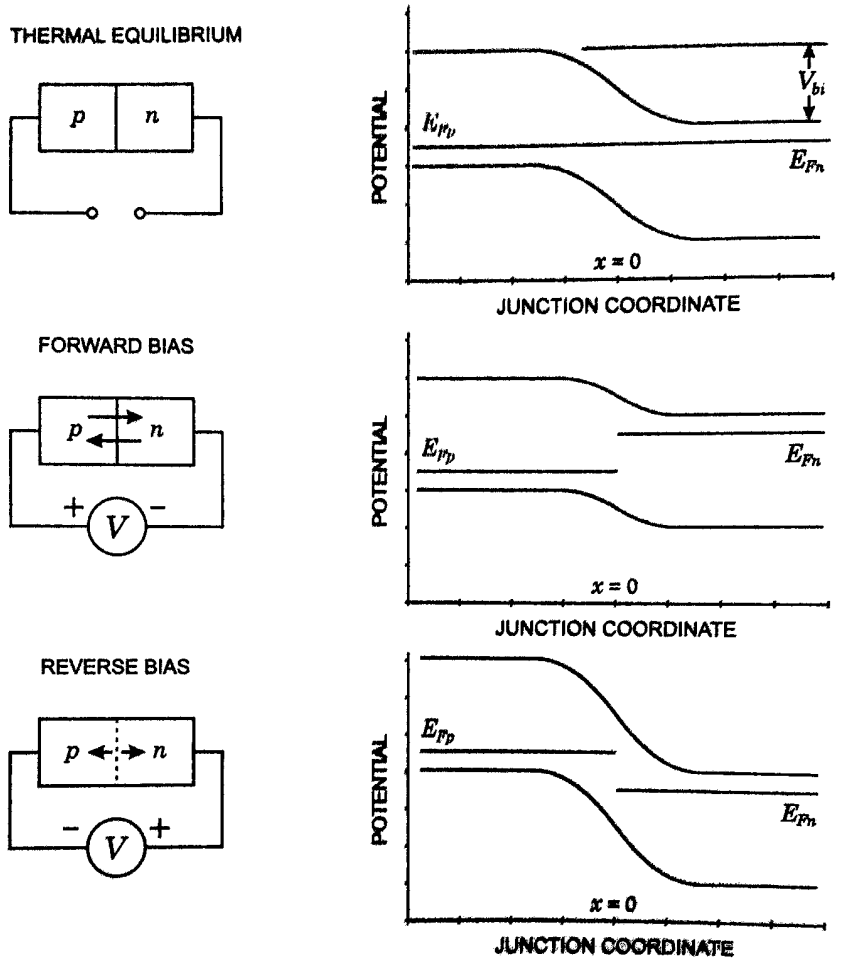


Figure 2.14: The effect of biasing on the potentials and Fermi levels on each side of a *pn*-junction from Spieler [2005], showing the potential levels and the Fermi levels on each side of the junction in each case.

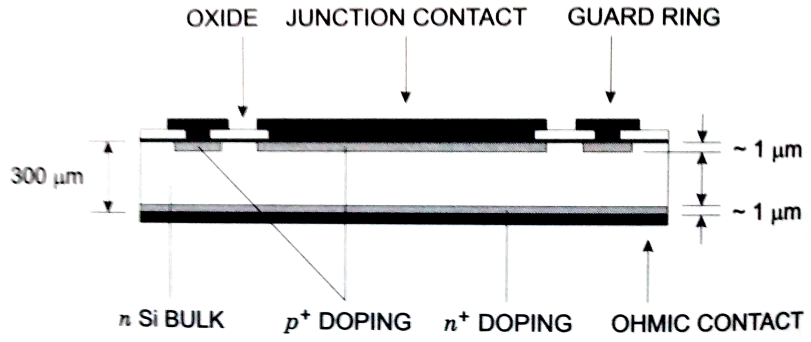


Figure 2.15: A detector structure from Spieler [2005] with the electrodes shown in black and the bulk shown between two heavily-doped layers.

a capacitor with the depletion region as the dielectric. The electric field in this region sweeps any generated electrons and holes to the electrodes where they can be detected and analysed.

If a detector is thin enough and the voltage across it is high enough, the depletion depth may extend almost the whole thickness of the detector, and it is described as fully depleted. This is particularly useful for a telescope design of instrument, where particles which transverse the front detector and also trigger another detector behind can be both identified and measured. This design concept is discussed further in section 3.2.2.

Detector structure

Detector diodes usually have a highly-doped layer at the surface and a lightly-doped bulk, with the depletion region reaching across the bulk. This structure is shown in figure 2.15 with the electrodes shown in black.

A reverse biased diode creates a depletion region with a width which depends on the bias voltage and the dopant concentration on either side of the junction.

The dead layer, or entrance window, is the initial thickness of the detector ma-

terial in which deposited energy cannot be measured by the system. This could include the electrodes or an oxidised layer, and must be included in simulations, since a large fraction of low energy particles will not reach through the dead layer to the active volume.

Models of ion energy loss

When a heavy ionising particle is incident on a detector, it will lose energy both through interactions with nuclei and electrons in the solid and through ionisation. The rate at which this energy is deposited in the detector is the linear stopping power $S = -\frac{dE}{dx}$, also known as specific energy loss, where E is the energy of the particle, and x is the distance travelled in the material. Since an ion is of a comparable mass to the nuclei in the solid, initial interactions will have relatively little effect on the direction or momentum of the ion; it is when the ion reaches lower energies that it may take a less well-defined path and lose the majority of its energy within the material at a relatively well-defined depth. This effect is more pronounced for heavier ions; the range of peak energy deposition becomes narrower.

This stopping power is well-modelled by the Bethe equation from Bethe [1930] which is expressed by Knoll [2000] as:

$$-\frac{dE}{dx} = \frac{4\pi e^4 z^2}{m_0 v^2} NB, \tag{2.3}$$

where

$$B \equiv Z\left[\ln\frac{2m_0v^2}{I} - \ln\left(1-\beta^2\right) - \beta^2\right]$$

and

$$
\begin{aligned}
e &= \text{charge of an electron} \\
ze &= \text{charge of the primary particle} \\
N &= \text{number density of the absorber} \\
Z &= \text{atomic number of the absorber} \\
m_0 &= \text{electron rest mass} \\
v &= \text{incident ion velocity and } \beta = \text{v/c} \\
I &\text{ is a constant of the material.}
\end{aligned}
$$

The data tables for stopping power and range of protons and alpha particles in solids are provided by the National Institute for Standards and Technology (NIST), which are produced using the Bethe formula with a density effect correction based on experimental results and analysis from Sternheimer [1956] and Sternheimer et al. [1982]. They are shown for silicon in figure 2.16 and figure 2.17. It can be seen that the heavier alpha particles are stopped more quickly by the silicon and have a shorter range. Heavier ions would show this effect even more strongly.

The stopping time is estimated by Knoll [2000] to be

$$T \simeq 1.2 \times 10^{-7} R \sqrt{\frac{m_A}{E}} \tag{2.4}$$

where

$$
\begin{aligned}
R &= \text{range in metres} \\
M_A &= \text{particle mass in atomic mass units} \\
E &= \text{particle energy in MeV.}
\end{aligned}
$$

This gives a stopping time and therefore a pulse length of a few picoseconds in a solid detector.

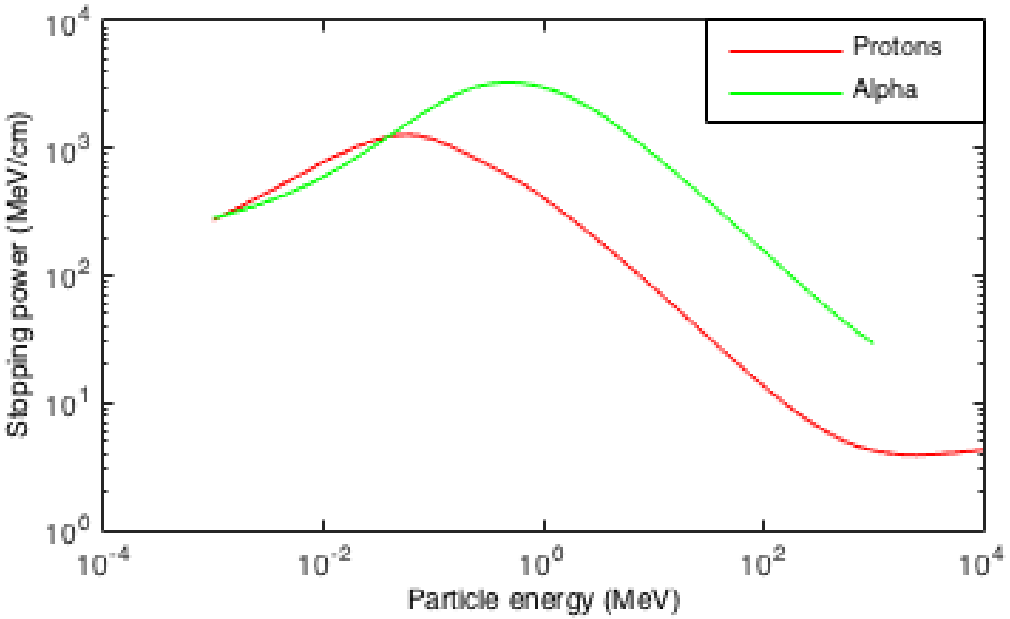


Figure 2.16: The stopping power of protons and alpha particles in silicon from the PSAR and aSTAR tables in the NIST databases (Available at http://physics.nist.gov)

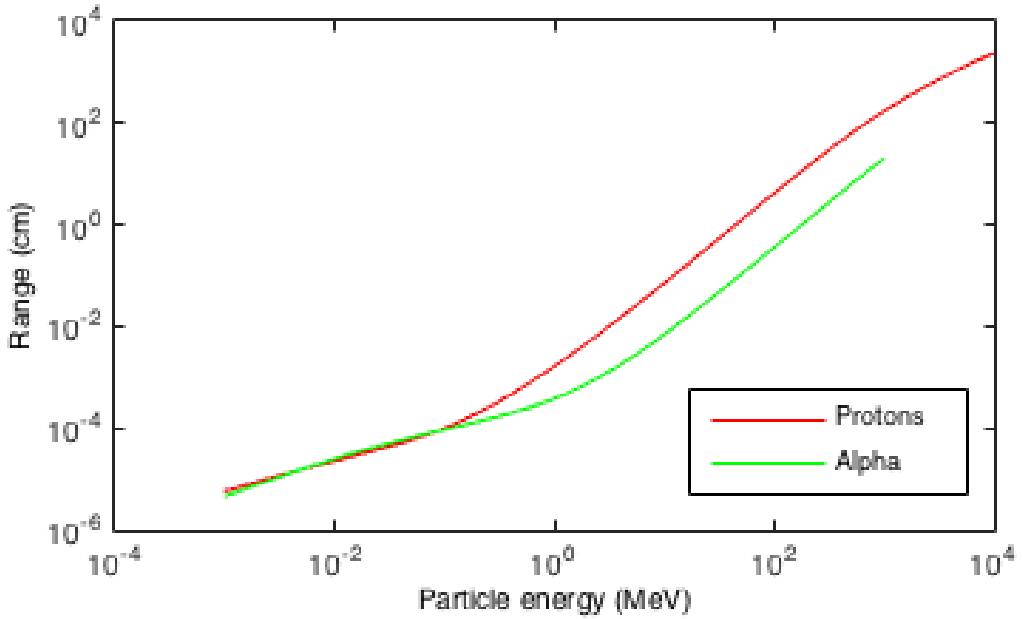


Figure 2.17: The estimated ranges of protons and alpha particles in silicon from the PSAR and aSTAR tables in the NIST databases (Available at http://physics.nist.gov)

Models of electron energy loss

The Bethe equation described above is used for almost all heavy charged particles. Electrons, being far lighter, undergo different behaviour in semiconductors - interactions with electrons in the crystal are capable of changing their direction significantly and greater changes in velocity create energy losses through bremsstrahlung at higher energies.

As a result of these processes, Bethe's adapted model from Bethe [1932] for the specific energy loss of electrons is given by a combination of a term for each; the expression for the specific energy loss due to ionisation and excitation (the "collisional losses") is expressed by Knoll [2000] as:

$$-\left(\frac{dE}{dx}\right)_c = \frac{2\pi e^4 NZ}{m_0 v^2}\left(\ln\frac{m_0 v^2 E}{2I^2\left(1-\beta^2\right)} - (\ln 2)\left(2\sqrt{1-\beta^2}-1+\beta^2\right) + \left(1-\beta^2\right) + \frac{1}{8}\left(1-\sqrt{1-\beta^2}\right)^2\right) \tag{2.5}$$

and the radiation, or bremsstrahlung, losses by:

$$-\left(\frac{dE}{dx}\right)_r = \frac{NEZ\left(Z+1\right)e^4}{137 m_0^2 c^4}\left(4\ln\frac{2E}{m_0 c^2} - \frac{4}{3}\right) \tag{2.6}$$

where the values are defined as in equation 2.3.

The data tables for stopping power and range of electrons in solids are provided by NIST; the stopping power in silicon is shown in figure 2.18 and the range in figure 2.19. The stopping power values are based on the Bethe formulae, with Sternheimer's corrections as described for ions and a mean excitation energy derived analytically, while the radiative stopping powers are evaluated from theoretical bremsstrahlung cross sections described by Seltzer and Berger [1985]. Analytical formulas using a high-energy approximation are used above 50 MeV, and accurate numerical results of Pratt et al. [1977] below 2 MeV. Values between 2 MeV to

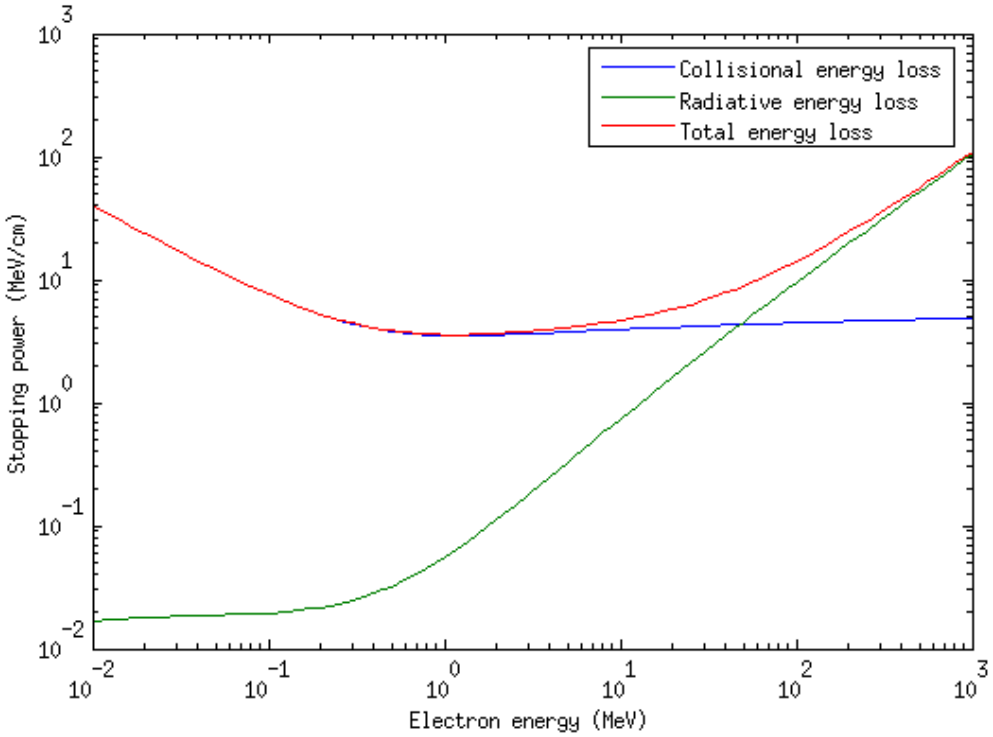


Figure 2.18: The collisional and radiative stopping power of electrons in silicon from ESTAR (Available at http://physics.nist.gov)

50 MeV are obtained by interpolation.

The collisional term dominates heavily for electrons up to several MeV, where the radiative term becomes significant and bremsstrahlung becomes an important consideration. For higher-Z materials, such as tungsten, the radiative term becomes significant at lower electron energies of several hundred keV.

Signal distribution in silicon particle detectors

The latest parametrisation of energy loss distribution through pixellated detectors, which are described in section 2.3.2, has been estimated analytically by Siklér [2012] and Siklér and Szeles [2012] for use in particle identification in large-scale high-energy

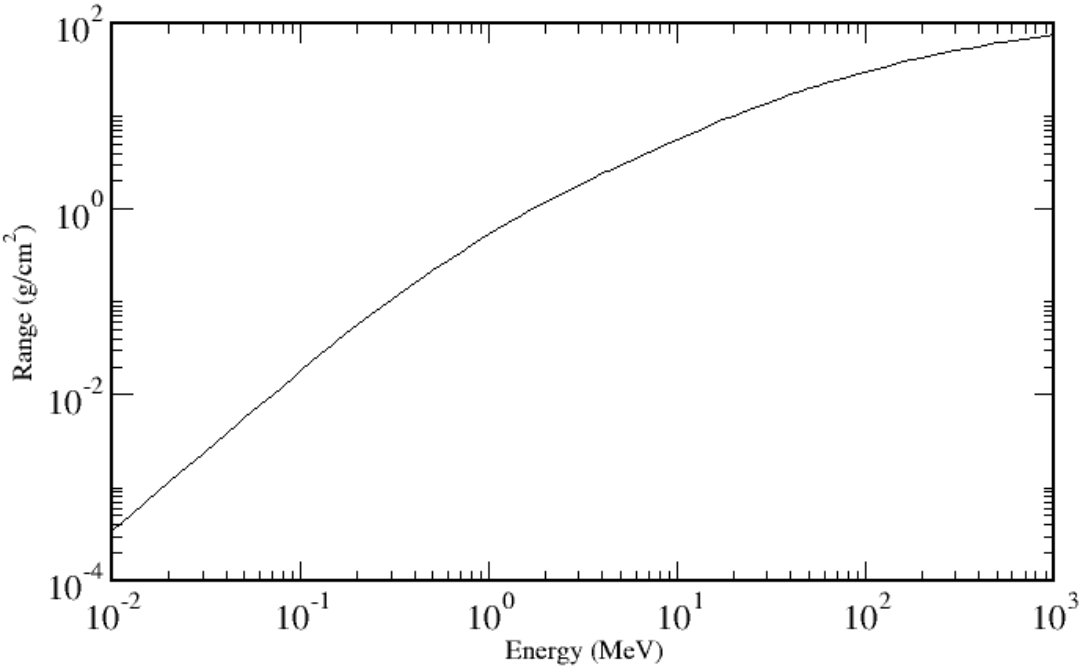


Figure 2.19: The estimated range of electrons in silicon (Z=14, N=$5.0 \times 10^{22}\,\mathrm{cm}^{-3}$ for silicon at room temperature from ESTAR (Available at http://physics.nist.gov)

accelerators.

Light charged particles, particularly electrons, lose energy both through collisional processes and through radiative processes. In contrast to heavy particle interactions, such as incident proton or alpha particle radiation, the low mass of electrons means interactions with electrons or nuclei within a solid medium can significantly change their path. Therefore the particle can change direction entirely and an electron may turn back on itself and leave the detector before it has deposited all of its energy. It may also deposit a more unpredictable amount of energy in the dead layer of the detector. This means an important part of such a detector design is involved with simulating and testing, so that the amount of energy deposited in the detector can be related to a probable flux, energy spectrum and direction of incident electrons. The validity of using solid state detectors to measure energetic electrons has been discussed by Vampola [1998].

Thermal noise

Semiconductor detectors are susceptible to thermal noise or 'dark current', when thermal motion produces a signal. Although thermal energies even at room temperature are around 26 meV, much smaller than the bandgap energy, random fluctuations can displace electrons in the lattice in the absence of an external signal.

The current generated by thermally-produced electron-hole pairs is a combination of the currents generated in each volume of the detector, which have slightly different temperature dependencies. Each component can be characterised by a form of Arrhenius equation, as described by Widenhorn et al. [2002].

The diffusion dark current is created at the field-free region, and is proportional to

$$I(T) \propto T^3 exp\left(-\frac{E_g}{kT}\right) \tag{2.7}$$

where

$$T = \text{absolute temperature}$$
$$E_g = \text{bandgap energy}$$
$$k = \text{Boltzmann constant}$$

and the constant of proportionality is related to the detector material, design and geometry, including the diffusion length. In the case of back-illuminated CCDs, this constant is related to the size of the field-free region rather than the diffusion length.

The second component, the depletion dark current, is caused by the thermal generation of charge carriers in the depletion region and is proportional to

$$I(T) \propto T^{3/2} exp\left(-\frac{E_g}{2kT}\right) \tag{2.8}$$

where the terms are defined as in equation 2.7, with a related, but different, constant of proportionality.

Which of these two terms - diffusion or depletion dark current - dominates depends on the structure of the detector and the temperature at which it is operating.

This dark current manifests itself in the output signal as a superposition of a constant offset and a white noise component. This can be significantly suppressed by running the detector at low temperatures, within the operating limits of the attached electronics.

2.3.2 Position-sensitive semiconductor detectors

A subdivision of charged particle detectors based on semiconductor devices includes those designed in such a way that the position or the angle of the incoming particles can be quantified.

This family of detectors has a wide range of uses in space instrumentation for

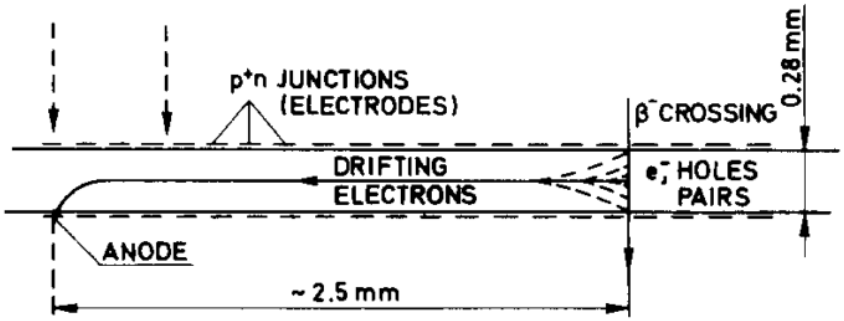


Figure 2.20: Diagram of the drift detector demonstrated in the original paper by Gatti et al. [1984]

measuring all wavelengths of electromagnetic radiation and many types of particle. In addition, their many uses on Earth, such as in cameras and high energy physics, has meant that the technology available is significantly advanced and tested.

Semiconductor drift detectors

The drift detector works by creating a potential well in the central layer of a detector, through the use of two reverse-biased semiconductor junctions and in addition a field parallel to the surface. The signal charge is collected on an anode at the end of the detector; the time taken to drift through the detector can be used as an accurate measure of the distance of the ionisation from the anode. Silicon drift detectors were proposed by Gatti and Rehak [1984] and demonstrated by Gatti et al. [1984]. A diagram of this original device is shown in figure 2.20.

With a known source, where the time that the particle or photon arrives at the detector is known, the drift time can be calculated easily. However, when measuring energetic particles in space, the time of arrival of the incoming radiation must also be measured and recorded, adding additional complexity.

The low capacitance possible with such devices means they can be developed to produce very low electronic noise, and are commonly used for X-ray instruments. For example, such a design is included in one of the instruments on the proposed ESA mission LOFT which is described in section 3.5.5.

Pixel detectors

The pixel detector is the simplest way of subdividing a detector into several sensitive areas - an array of small detectors positioned together can separate the signal from incident radiation at different areas on the detector. A particle instrument might use this type of detector with an aperture placed in front of it to break down flux measurements into multiple angular bins. The resolution that can be achieved with such a detector is limited, but pixel sizes down to tens of micrometres are now possible.

The STEIN instrument described in section 2.2.2 used a four-pixel detector with an electrostatic collimator to measure species, angle and energy of energetic particles.

Microstrip detectors

A microstrip detector has one electrode subdivided into strips, so that the point where the signal is created can be measured in one dimension.

The strip pitch, and therefore resolution, is typically tens of micrometres, but the length of the strip - the width of the detector - is unlimited.

An extension of this design uses subdivided electrodes on each side at an angle to each other to estimate the signal position in two dimensions. Both the simple one-dimensional microstrip design and the two-dimensional design with electrode strips on both sides are shown in figure 2.21. In the two-dimensional case, the electrodes receiving the signal are highlighted, showing how the location of the incident radiation can be recovered.

This two-dimensional microstrip design becomes ambiguous when multiple hits occur at the same time. To mitigate this problem, it is possible to have the two

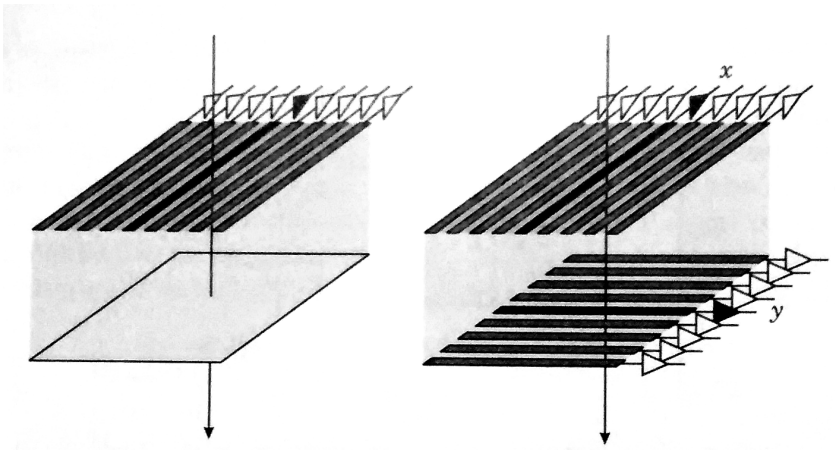


Figure 2.21: Microstrip detector electrode designs from Spieler [2012]

sets of electrodes a small angle offset from each other, rather than perpendicular. This reduces the number of intersecting electrodes when the detector is viewed from above, so the chance of ambiguous events occurring is reduced.

Charge-coupled devices

The original concept for the CCD was developed by Boyle and Smith [1970] and demonstrated by Amelio et al. [1970] - a development for which Boyle and Smith were awarded the 2009 Nobel prize in physics.

Using a series of electrodes above each pixel, charge collected over an integration time in each pixel can be moved through the detector, one row at a time, to a serial output register and in turn the register can be read out one pixel at a time, as shown in the scheme in figure 2.22.

Since their invention in the 1970s, CCDs became very popular and were used in almost all digital cameras. Consequently, the size, speed and resolution have improved greatly through further developments in semiconductor manufacturing; the improvement in pixel size and therefore resolution of CCDs is shown in figure 2.23.

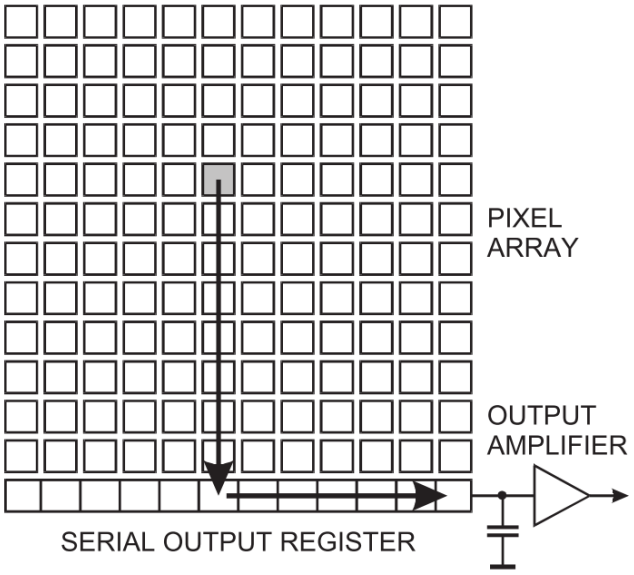


Figure 2.22: CCD output process from Spieler [2012]. Each row is shifted in turn into the output register whence each pixel is read out in turn.

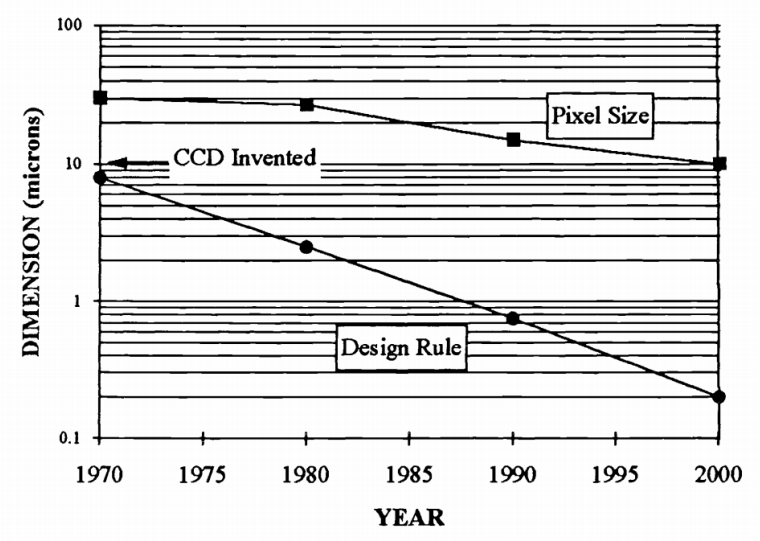


Figure 2.23: Development in the size of CCD pixels from Fossum [1993]

In recent years, CCDs have largely been replaced by complementary metal-oxide-semiconductor (CMOS) detectors in many applications, but are still very important in several other fields, for example X-ray astronomy. It is one of these X-ray focused CCDs that has been used for the experiments described in chapter 5.

Monolithic active pixel sensors

In an active pixel sensor imager the charge is converted to charge in each pixel individually before being read out; the principle of individual pixel signal amplification was described by Noble [1968]. Since at the time, CCD technology was more reliable, CCDs dominated consumer technological applications. As higher-stability CMOS sensors were developed in the 1980s and 1990s, active pixel sensors have come into far more general use, as predicted by Fossum [1993] in "Are CCDs dinosaurs?". For scientific applications, each technology has its merits. CCDs can have very low noise characteristics and are often the detector of choice, for example in X-ray imaging. However, active pixel sensors can be far more flexible in readout schemes.

Typically, the electronics for each pixel readout are situated next to or behind the sensor of each pixel. The layout depends partly on the type of expected radiation and whether it is likely to cause damage to the readout electronics. Monolithic CMOS detectors have been developed for charged particle tracking for example by Deptuch et al. [2003] and are commonly used in high energy physics.

2.3.3 Radiation damage

At the beginning of its life, the crystalline structure of a detector is close to perfect. Once in a radiation environment, the lattice begins to suffer from unpredictable disruptions caused by incident particles and photons. The two mechanisms by which radiation damage occurs in semiconductors are displacement damage and ionisation damage; overviews by Knoll [2000] and Spieler [2005] have described these mechanisms.

Displacement damage occurs when silicon atoms are displaced from their lattice sites and the resulting crystal defects remain and change the electrical characteristics of the detector. For this to occur, incident radiation must deposit sufficient energy in one lattice site.

Any damage to the crystal structure of the detector bulk will create trapping sites for charge carriers, and cause an increase in leakage current. While particles below around 25 eV are not usually a concern, since they do not carry sufficient energy to alter the crystal's structure, higher energy photons and electrons as well as even low-energy protons and ions will eventually degrade the quality of readings produced by a detector. Different particle types interact differently with the silicon lattice, as shown in figure 2.24, where the effect of radiation has been scaled to the effect of 1 MeV neutrons.

Ionisation damage occurs when ionisation is triggered in the oxide layer and charge carriers are liberated and then trapped elsewhere in the detector. Radiation damage to the silicon dioxide layer of the detector behaves differently to damage to the detector bulk due to the different, highly irregular, crystalline structure. This

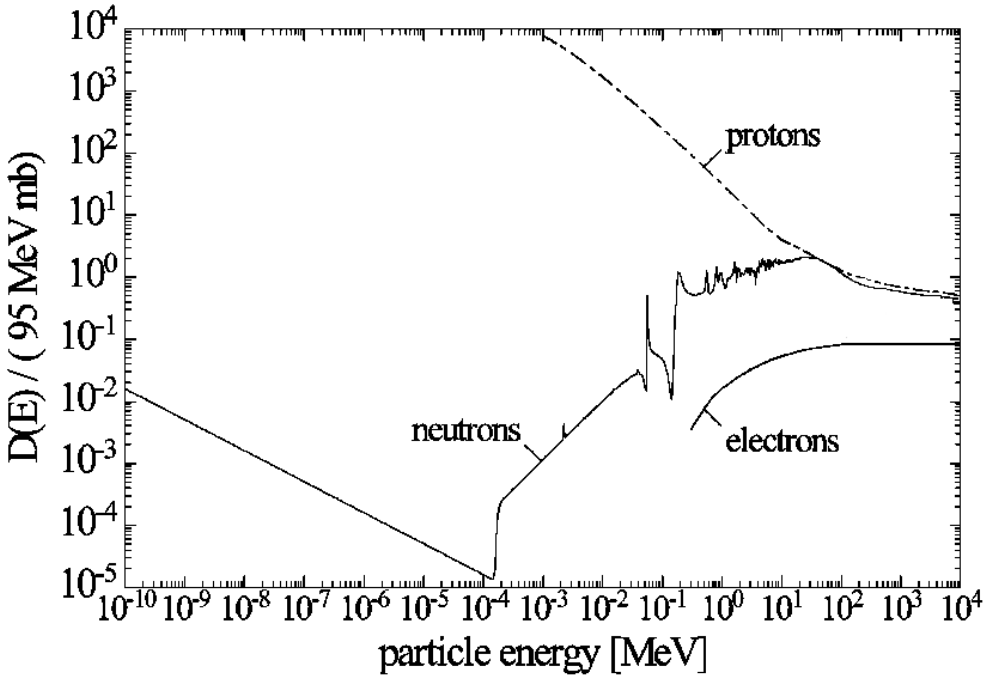


Figure 2.24: Displacement damage functions for neutrons, protons and electrons in silicon relative to damage from 1 MeV neutrons from Lindstroem et al. [1999]

structure is more tolerant to radiation damage, but ionising radiation induces charge carriers in the surface layer. Holes diffuse away more slowly than electrons, and this disparity changes the voltage distribution in the device.

Radiation damage will tend to degrade the detector performance over time since the increase in leakage current reduces the signal to noise ratio. At low damage levels, this manifests itself in an increase in reverse bias current proportional to the temperature, as in the case of thermal noise. This leakage current may eventually lead to catastrophic sensor failure in the event of a thermal runaway. However, many detectors can be nearly restored to their original characteristics through annealing, which reduces the leakage current by removing the structure of radiation-induced defects. This is a process of heating the detector to a temperature of the order of 100 °C for long enough to allow the lattice structure to re-establish itself through thermal fluctuations.

Silicon detectors can be designed for radiation hardness, both by changing the geometry and altering the density of impurities in the silicon bulk. Oxygen and carbon are the main impurities which affect radiation hardness. Different types of radiation create different ratios of *n*- and *p*- type defects depending on the impurities of the silicon; this can be used to compensate for radiation damage to detectors in mixed radiation environments.

For a particular detector in a particular environment, both the total expected radiation dose and the expected flux and type of higher-energy particles must be known to understand the expected lifetime of the detector and the shielding needed for the instrument. For this purpose, integrated tools have been developed, including Geant4 Radiation Analysis for Space (GRAS) which uses Geant4-style 3-dimensional geometries with common types of space radiation analysis.

The integrated web-based software tools available from the Space Environment Information System (SPENVIS) are based around similar Geant4-based tools and are designed to make the process of estimating the radiation environment simpler. For example, it is straightforward to calculate the flux of protons above a certain

energy in a sensitive volume with planned shielding for an LEO with a specified altitude and inclination. These calculations are based on the AE-9 and AP-9 models of electron and proton fluxes in the magnetosphere.

This estimated dose and flux must then be used to estimate the degradation and lifetime of the detector through its mission, and to make a plan for any annealing required. Although this can be partially predicted through such simulations, in general particle detectors for flight will undergo radiation testing to assess their ability to operate in space, as with any other component.

2.4 Semiconductor detector electronics design

In a semiconductor radiation detector device, the electron-hole pairs formed when a sufficient amount of radiation energy is deposited in the sensor are the charge carriers forming a small current pulse across what can be treated as a capacitor in a circuit design.

The current pulse from the detector needs to undergo analogue transformation before it can be counted or measured and recorded by the digital part of the system.

2.4.1 Detector operation modes

Generally, a radiation detector can operate in current mode or pulse mode.

A detector in current mode measures the time-averaged current produced across the terminals to gather information about the product of the interaction rate and the charge per interaction, and tends to be used when event rates are high.

Pulse mode is capable of providing information on the timing and amplitude of individual events - this is the readout mode which is used when the energy or timing information of each event is required. In a plasma analysing instrument, as opposed to a dosimetry instrument, each current pulse caused by an event would be read out and likely quantified.

2.4.2 Elements of detector electronics design

A possible, and common, layout for pulse readout circuitry would involve a preamplifier and pulse shaper before the digitiser.

Sensor

In a semiconductor detector, the energy of the incident radiation is deposited in the sensor. The charge formed on the electrodes is related to the voltage across the semiconductor.

Amplification

A charge-sensitive preamplifier, or active integrator, amplifies the charge on the detector and converts it to a voltage.

Since the energy deposited in the sensor by the particle or particles is very small, the signal charge it creates is correspondingly small, rarely more than a few fC (Spieler [2005]), so the signal must be amplified. Even tiny noise levels or fluctuations before this stage will cause large offsets in the final data, so the design of the system must take care to minimise potential electronic noise sources, for example defects in the detector or long or unshielded harnessing.

Filtering

The pulse shaper element of the system is designed to reduce the noise levels as far as possible by applying a filter that will favour the signal above the noise.

By predicting the frequency spectra of the expected signals, the relevant frequency band or bands can be highlighted above the broader-spectrum noise that originates both in the detector (as dark noise) and in the preamplifier electronics.

Digitiser

The analogue components must be converted into a digital signal for computational processing. There are multiple ways of doing this.

An analogue-to-digital converter (ADC) will digitise the shape of the pulse. Since the pulse length is shorter than the speed of almost all circuitry will permit, the pulse has to be amplified and extended - the height of the pulse will at least be enough to measure the deposited energy of the event, and more sophisticated analogue electronics are capable of also identifying some information on interaction position and path.

A pulse height analyser will hold a voltage at the peak height of a pulse to be measured. Since the height of the pulse is proportional to the energy deposited in the detector, this will give a measure of the energy of an event, but since the measure is of only one point on the curve, noise will add errors to both the time and energy data of the interaction.

A counter, or discriminator, will count pulses with a peak above a certain height. Several can be used, set at different discriminator levels to gain a low-resolution spectrum of the deposited energy spectrum.

Chapter 3

Instrument concept

The motivation for the design and development was to produce an energetic particle detector suitable for a CubeSat or small satellite platform to measure particles in the energy range of 10s to 100s of keV.

Using a solid state detector that is capable of relatively low-energy particle measurements is one way of reducing the required instrument resources in the energy range of 10s to 100s of keV - this has more often been done with electrostatic analysers in the past. Recent developments in silicon detectors for space allow very thin entrance windows to be manufactured. This can allow the detection of individual electrons as low as 2 keV, and protons in a slightly higher energy range. Using a solid state detector for this particle energy range rather than MCPs or electrostatic devices can reduce both the volume required and the voltages needed.

The CATS instrument was designed, characterised and built into a demonstration flight instrument at MSSL as described in the PhD thesis by Bedington [2012]. While highly capable at identifying particle energies and angles in high flux environments, its use in low flux space environments was limited by its low geometric factor and very small field of view. This highlighted a gap in the capabilities of current miniaturised particle instrumentation that this project aimed to fill: the ability to measure accurately the angular distribution of suprathermal charged particles in low

flux environments.

The choice of design concept was driven by the very strict engineering requirements for instruments on small satellites, primarily of power, as well as a short timescale for design, manufacturing, testing and documentation, all of which would be common on CubeSats in particular. The resulting design would then be suitable for any mission where simplicity and miniaturisation are high priorities. Such an instrument is then also potentially adaptable as a small radiation detector for other, similar environments, depending on the flexibility of the design.

Further detailed requirements are given in the following sections.

3.1 Scientific objectives and requirements

The aims of the instrument to be designed are to gain detail for modelling to help understand space weather behaviour or improve space weather predictions. Enhancing these data would require an improvement, for example in spectral range or resolution, angular range or resolution or time resolution, or in utilising the multi-position measurement possibilities of swarm missions.

As described in table 1.1 in section 1.3, hazardous particle populations between 1 keV and 1 MeV in the LEO environments of the auroral regions and the SAA vary over timescales of minutes or possibly hours. Tracking these changes would therefore require a measurement at least once a minute.

In addition, the spin rate or pointing stability of the spacecraft platform, while providing the opportunity to sample the particle environments through a range of angles, may also be a limiting factor of the exposure time or cadence of a measurement.

Directional particle populations may also be of interest; as described in section 1.3.4, radiation belt particle flows may have directional fluxes of the order of 5° and other space environments defined by magnetic fields will also have aligned particle fluxes, such as electron strahl and magnetised planetary environments.

With these possibilities in mind, the purposes of the design and use of this instrument would be:

1. Detection and measurement of energetic particles in a radiation belt environment
2. Demonstration of a novel highly-miniaturised silicon detector based system.

These objectives could be optimised or adapted for other environments with similar plasma densities and angular distribution, for example in the high-radiation environments near Jupiter's poles.

3.1.1 Requirements for the design

Although there is significant variation in the requirements placed upon a CubeSat instrument, in general, the space and power remaining for a science unit after the systems for telemetry, attitude control and power and control subsystems have been implemented is usually restrictive. The requirements described here might be typical for an instrument on a 2U CubeSat that has no or minimal other experiments on board. They will be taken fairly loosely when considering potential designs since CubeSat sizes and power supplies vary but they must be borne in mind to limit possibilities to potentially practical ones.

Performance requirements

In order to meet the objectives, some or all of these aims might be considered.

Capable of detecting electrons, protons and ions within the spacecraft's operating trajectory
Can discriminate between types of particle (ideally electrons/protons/ions)
Can identify the angular distribution of particle velocities with a precision of at least 10°
Can recover an energy spectrum in an energy range from 1 keV to 1 MeV

Requirements for the system

These requirements aim to take into account the qualities the design would need to meet the objectives given above in the environment of a CubeSat, including those related to signal-to-noise ratio (SNR), power and radiation tolerance.

Mass less than approximately 200 g
Power less than approximately 500 mW
Stability over fluctuations in temperature from −30 °C to 20 °C
Capable of operation and measurement for 1 kRad total ionising dose

The system requirements above lead to requirements on each of the components of the instrument.

Requirements for the detector

The detectors considered for an energetic particle detector are usually silicon- or semiconductor-based, since an MCP has limited energy and position resolution and reduced efficiency above tens of keV. In addition, solid state detectors have some strong advantages in miniaturised instruments, as described in section 2.2.1.

Those considered in this section should meet the minimum requirements to be suitable for flight on a CubeSat:

Reliable operation over fluctuating temperatures, with no strong thermal control requirements
SNR set to the level needed for particle measurement with detector and readout mode used
Radiation resistant or shielded enough to remain functional for the duration of the mission
Voltage requirements no more than the order of 100 V

Since any detector will need associated electronics, these will also have closely-

related requirements:

Low power (below approximately 500 mW)
Low noise (comparable to detector noise)
Capability to produce required voltage across the detector
Ability to withstand radiation with shielding for a mission duration of three months

Requirements for additional optics and geometry

When considering geometry and materials, realistic design concepts must be effective in their purpose and practical to manufacture and fly.

Providing effective shielding of particles not of interest with $0.24\,\mathrm{g/cm^2}$
Stability against UV and particle radiation for 6 months under direct exposure to sunlight
Stability over fluctuations in temperature from $-30\,^{\circ}\mathrm{C}$ to $20\,^{\circ}\mathrm{C}$
High physical strength against vibrations and pressure differences as specified in the CubeSat standard
Very low outgassing levels

3.2 Design trade-off

The very strict constraints demanded by the requirements of the CubeSat platform necessitate a very much reduced range of design options as compared to large-scale missions.

This section will cover thoughts on possible detectors and geometries that were considered during the design process.

3.2.1 Detectors

For these applications, the possible detectors were somewhat limited. However, if the resource budget were flexible, the scientific returns would rapidly increase with detector capability.

A single detector

For the most minimal power requirements, a single detector would require only one set out of readout electronics and therefore approximately half the resources of any two-detector setup. However, clearly the data that could be gathered from such a setup would be limited.

A discriminator in the readout electronics could be chosen to select the cut-off energy for particle detection, for the lowest-resource option, or a slightly more complicated digitisation step might allow a good readout of the particle energy spectrum. However, there would be no simple way of recovering particle species or direction information.

Position-sensitive detectors

There are several types of more complex detectors which are capable of retrieving further information, such as the position of an impact on the detector, some of which are described above. The power requirements are relatively high for a CubeSat, although some may be able to manage certain types of detector. CCDs and CMOS detectors are commonplace and have often been used in cameras on CubeSats.

Pixellated detectors such as the four-pixel detector produced for STEIN would be possible. Glaser et al. [2009] described the detector readout electronics as being 90 g and 290 mW in total, and the 150 V voltage supply as 75 g and 250 mW - if this power could not be supported continuously on a CubeSat, an instrument could be run for only part of an orbit.

Medipix

The Medipix sensor family is a series of CMOS detectors developed at European Organization for Nuclear Research (CERN) (Amendolia et al. [1999]), designed as a prototype CMOS imaging chip capable of counting photons. Originally the first generation, which was received in 1997, was 64×64 pixels of 170 µm square, while the second- and third-generation versions have a higher resolution of up to 55 µm. The Timepix chip is very similar to the Medipix chips, but with more readout options, including the ability to measure the point in time at which a particle reaches the surface.

They have found a wide range of applications, for example in radiography, and are of interest to the field of space instrumentation - for plasma instrumentation, but also cosmic rays, dosimetry and neutron detection.

Medipix and Timepix sensors were first launched in 2012, both on a NASA mission for radiation detection on the ISS, and in the Langton Ultimate Cosmic ray Intensity Detector (LUCID) on TechDemoSat-1. These mission plans have been described by Pinsky et al. [2012] with the aim of producing data to form the basis of the development of a new generation of photon counting detectors designed primarily for space. Since the launch, the performance of these missions has been analysed by Pinsky et al. [2014] and Whyntie and Harrison [2015].

The complexity of the device means the resource budgets may be stretched, especially when including the readout processing which will vary with the complexity of the data needed. Llopart et al. [2001] give a figure of approximately 500 mW as the power required by the analogue electronics, and the digital electronics could be designed to be significantly lower power than this. In addition, the Medipix system can be made very light and compact, as in the Universal Serial Bus (USB)-run examples on the ISS from Pinsky et al. [2014].

RADFET

A RADFET could be used for dose measurements. This alone would be far less useful than also measuring the energy spectrum of the particles, but could be used if the requirements could not be met otherwise, or in conjunction with another sensor. Dosimetry measurements are particularly of interest for estimating radiation dose to astronauts.

Previous missions have flown small RADFET-based plasma instruments, including MuREM and SSTL RM as described by Taylor et al. [2012]. These instruments have varying power requirements, for example MuREM has a low power mode of 0.3 W, so a RADFET alone would be significantly lower.

3.2.2 Geometries

Two adjacent detectors

The CubeSat Charged Particle Detector described by Dowler et al. [2002] used two small ORTEC detectors facing in the same direction with thin entrance windows and depleted regions. One of these detectors would have a shielding foil obscuring it, of a thickness which would give the two detectors different and adjacent energy ranges. Although this design did not launch, it would be the simplest possible set up with two detectors and would be able to give additional data beyond one detector - an improvement in the range of the energy spectrum.

In addition to this proposed design, a similar concept is used as part of SREM, although one of the detectors has a further detector behind it.

Telescope design

In order to obtain the most information about the mid-energy-range particles, a telescope arrangement allows a degree of particle discrimination. While the spectrum of particle energies can be found for the lower energies, the particles that pass through the first detector and deposit further energy in the second detector can, in

theory, be identified by the relevant output pulse heights from each of them, which are compared as the ΔE-E characteristic. In order to make this discrimination feasible, the noise and variation in the output pulse height data need to be very small, and the coincidence logic required to identify when a particle deposits energy in both detectors at almost exactly the same time adds significant complexity to the electronics required for the instrument.

The telescope design is a very common one for many types of particle detection and identification. Previous space plasma instruments to use a pair of detectors as a telescope include part of SREM and the Lunar Particles Shadows and Boundary Layer on the Apollo 15 and 16 subsatellites. Even more instruments have made use of larger detector stacks, especially on larger missions.

Some investigation into the simulation of a telescope concept is given in section 3.3.1.

Two perpendicular detectors

The potential benefits of having two detectors facing in perpendicular directions would include the ability to gain more angular information. Additionally, depending on the geometry, if coincidence logic were included, any particles passing through both detectors might be identified and any strong variations in particle direction could be found and investigated.

A further development of this might be three orthogonal detectors to give a wider range of angles. A similar concept was used with Medipix detectors as part of the LUCID instrument described by Whyntie and Harrison [2015].

The use of a foil

A material can be placed in front of the detector or detector system, in order to affect the spectrum of energies impacting on the detector. A thin sheet might cut out the lower energy incident particles from the impinging energy spectrum, while a slightly thicker foil is often used to remove incident protons and ions while detecting

electrons. In this case, since the aim is to detect at the least the lower possible end of the energy spectrum, at least part of the detector system should remain uncovered.

The shape of the foil might also be taken into consideration. The path of a particle impinging on such a foil will be diverted through interactions in the material, and the modelling of this behaviour will be described in chapter 4.

Since the size of the detector is unlikely to be large, given the relatively high-density plasma environment in low-Earth orbit, the foil would be expected to be small and should be unlikely to need additional mechanical support, although it will need to support its own weight.

Almost equivalently to a foil, a coating could be used on the detector itself, if the aim is to attenuate some or all incoming particles. This would allow a smaller thickness to be used and be more structurally sound. In the case of a pixellated detector it would be necessarily done this way to reduce particles being deflected to the wrong pixel.

The use of a mask

A mask, the concept described in section 3.5, would be a novel design for energetic particles, in particular for electron detection.

Skinner [2004] has given an overview of coded apertures and their limitations. In many situations, a mask is not the best choice, but in a relatively high-flux environment where a minimal resource design is the highest priority, such as in this case, the design is worth investigating.

On one hand, the geometrical advantages of a large view angle with a large aperture area can increase the geometric factor of the instrument with a very low resource requirement, while on the other hand, there is limited angular resolution in the case of very low flux environments: the angular discrimination is only possible statistically and not for individual particles.

3.2.3 Operation and readout electronics

A description of the energy spectrum could be obtained by cycling through a range of energy bands at a rate significantly faster than any rotation of the satellite and the variation in time (from the frame of the spacecraft) of the objects of interest - for a spinning satellite the time period of rotation might be predicted to be of the order of 10 s.

The choice of digitiser will make a large difference to the resource requirements of the instrument. A simple discriminator would be the lowest-resource, a pulse height analyser would give more information, but use more power, and a complete ADC capable of digitising the whole shape of a pulse would likely use resources beyond those available to the smallest sizes of instruments on a CubeSat.

3.3 Initial conclusions

Designs incorporating one or more of these ideas have been considered. The concepts considered to be the most interesting and feasible are described here and summarised in table 3.1.

The particle telescope is interesting due to the possibility for particle discrimination and spectrum determination, but is not able to measure the angular distribution of the incoming radiation. Mask-based methods would be able to determine this, but require a Medipix detector or advanced geometries for a pixellated detector to be able to measure particles type and spectrum.

A telescope could be realised with a reasonably low power budget - the quoted power consumption consists of the detector and charge sensitive pre-amplifier - but would likely require additional, possible high-voltage, biasing, which is not included. The data processing for the telescope would require more advanced electronics for the coincidence logic and pulse shape analysis.

Mask based concepts would require around 500 mW for a pixellated detector, and more for the Medipix. The power consumption given is based on the detector

Table 3.1: Concepts that have been considered

	Telescope	**Mask&pixellated detector**	**Mask&Medipix**
Particle discrimination	Yes $>\sim$1 MeV p^+	No *without back detector*	Yes
Energy spectrum	Yes *limited range*	Maybe *depends on detector*	Yes *coarse resolution*
Angular distribution	No	Yes *statistically*	Yes *statistically*
Estimated mass	200 g	250 g	250 g
Estimated power	100 mW	500 mW	$>$ 500 mW
HV required	Likely	No	No
Processing	Complex *Coincidence logic* *Pulse shape*	Medium *Deconvolution*	Medium *Deconvolution*

power consumption and does not include additional electronics required for the signal processing. High-voltage biasing is usually not required for CMOS detectors.

3.3.1 Basic telescope design

Using two very straightforward silicon detectors, a measure of ΔE-E could be found for a small range of energies, recovering particle species and energy. The limit in the possible energy range of the recoverable spectrum comes from the voltages required across a thick back detector. Thicker detectors need significantly higher voltages to maximise the active region, and on many CubeSats it will not be feasible to create these high voltages. One way around this might be to use two or more detectors in place of the one at the back, with both connected to the same voltage source and readout circuitry. However, this will still increase the mass of the instrument and complicate the geometry and range of accepted angles of heavy particles.

In the case of this design, the energy spectrum can therefore be controlled to some degree by the number and thickness of the detectors and by the coating or shielding in front of the first detector.

While the concept of a two-detector telescope is far from novel, the attempt to reduce the size and complexity of the instrument beyond anything previously flown would make it of interest for a wider range of small-scale missions.

Design considerations

The thinner the entrance detector, the lower energy at which particle types can be distinguished. The simulations below used detectors from ORTEC - a fully-depleted D-series detector 15 µm in front of an ULTRA detector with a deep depleted region.

However, with very thin entrance windows, these detectors are likely to produce a signal from light. With an instrument of this design, stray light would need to be eliminated and measurements in the direction of a light source could not be taken.

Telescope simulations

The simulation environment used in this section is described in more detail in section 4.1 which covers the Geant4 toolkit and the Multi-Layered Shielding Simulation Software (Mulassis) tool which makes use of it for planar geometries.

To simulate the smallest possible silicon particle telescope design, which could be used under the extreme mass and power requirements of, for example, UCLSat, Mulassis was first used with different thicknesses of detectors commercially available from ORTEC.

For the simplest attempt at running the simulations, a macro was created for each particle type, energy and angle. Then each of these was run multiple times, and the dose in each layer was recovered, along with the pulse height or fluence between layers. To find the result of a simulation of any environment, the results of these can be added together in the correct proportions of angle and energy as well as particle type (electron, proton and alpha).

The initial simulation was run with a very deep back detector with the aim of capturing all the energy deposited by all incident particles from 0 MeV to 5 MeV. While the proton and ion spectra were relatively very easy to identify at higher energies, electrons were found to be sufficiently unpredictable that recovering the electron energy spectrum would be expected to be very difficult and identification of particles would usually be unrealistic. Solutions might be found to reduce this problem, or the instrument could be designed to focus only on ions, for example by sweeping away electrons using a magnetic field.

As an improved setup, further simulations were run using a reduced depth of the active region of the back detector to allow many of the electrons to pass right through while stopping all protons up to approximately 6 MeV.

Figure 3.1 shows an example of the behaviour of parallel beams of particles evenly distributed between 0 MeV to 5 MeV in Mulassis through a simple theoretical setup: a 15 µm front silicon detector and a 300 µm back detector with 1.39 µm Al shielding and the commercially-produced entrance and exit windows: 80 nm at the entrance

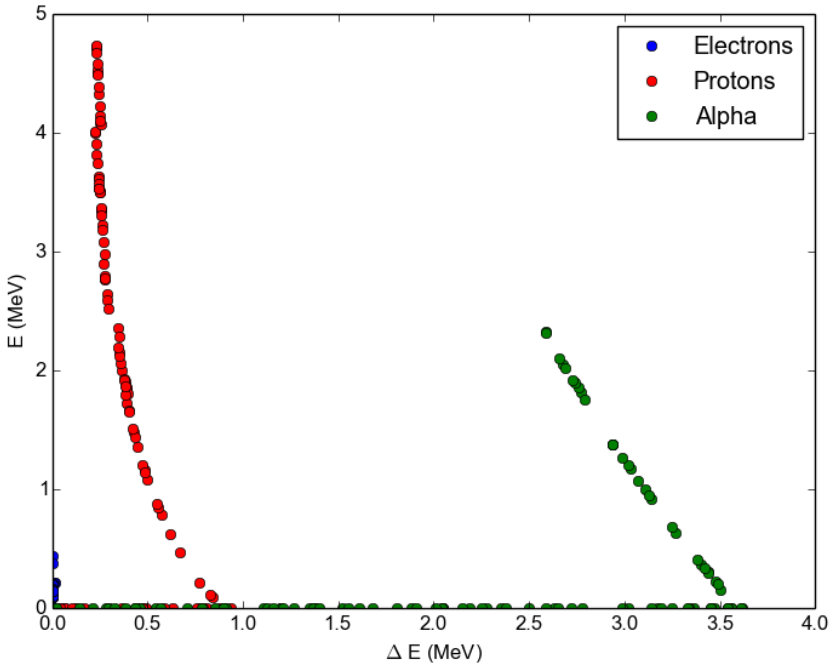


Figure 3.1: An example of ΔE-E for an even distribution of electron, proton and alpha particle energies from 0 MeV to 5 MeV where ΔE is the energy deposited in the front detector and E is the energy deposited in the back detector of particle. This shows that for this setup, particle identification between protons and alpha particles may be possible above about 1 MeV protons and 3.6 MeV alpha particles.

of the first and 50 nm at the second, and 225 nm at the front of the rear detector. The thickness of the active region of the back detector is enough to absorb almost all particles in the simulation that reach the front detector - only a few electrons would reach the back of it.

The very lowest energy protons do not even reach the active area of the front detector and produce no pulse in either detector readout. With increasing energy, the particles are shown along the bottom of the graph as all their energy is recorded in the first detector. The curve contains particles that deposit energy in the first detector and are stopped entirely by the second; with increasing energy the ratio of ΔE-E decreases - the highest energy protons are shown as points at the top left of the graph.

This is a somewhat idealised view, since the geometry is purely one-dimensional - the layers reach out to infinity. A real instrument would contain edges and collimation. If this design were to be taken further, a specific application for the design would need to be written in the Geant4 toolkit described in section 4.1.1. This would also be a far more efficient way of running this simulation, as it would avoid setting up the geometry repeatedly for each particle to be flown.

3.3.2 Mask and position-sensitive detector

Another concept that appears to be both feasible and interesting is that of a mask in front of a position-sensitive or pixellated detector. In the same way as with the telescope described above, shielding or preferably a coating on the detector would be needed to reduce the noise caused by low energy particles and radiation.

While the complexity of the readout electronics is relatively high, the support for these is likely to be available on the same board or chip as the detector and should not add much to the complexity of the design. The type of detector itself needs to be considered: a Medipix chip may be difficult to acquire but would have greater capability for directional sensitivity and particle identification, whereas a CCD or CMOS detector would be sufficient to gather statistical data on particle direction.

The mask, however, is a very low resource system that should be able to identify very narrow features in the energetic particle velocity distribution, such as the very flat linear pattern in the trapped particle environment, and may therefore be the right choice for a CubeSat instrument. This broad family of designs will be studied through simulation and experiment.

3.3.3 The decision process

Each of these concepts could be designed with a wide field of view and be produced on a size and power scale suitable for a CubeSat.

A telescope design has the potential to perform particle differentiation and produce a high energy resolution. It is, however, highly limited in ability to recover angular information.

The mask concept has the advantage of the possibility of retrieving a very high angular resolution for the distribution of particles, but with little ability to provide information on individual particles.

The mask concept was judged to have sufficient novelty, being the first proposed use of the technique for particle detection in space. Development of this would demonstrate a new concept and build on the high-resolution instrumentation already developed at MSSL. Because of this reason and the capabilities such an instrument promised to be able to show, beyond other miniaturised instruments, it was decided to investigate the coded aperture concept for this study.

3.4 Proposed design

The proposed design is driven by the very strict requirements - a mass below 200 g and power below 500 mW, as well as an optimised geometric factor. This leads towards a design with as few elements as possible.

The proposed design of a pixellated detector behind a mask is considered to be both novel and practical; a diagram of the elements is shown in figure 3.2. However,

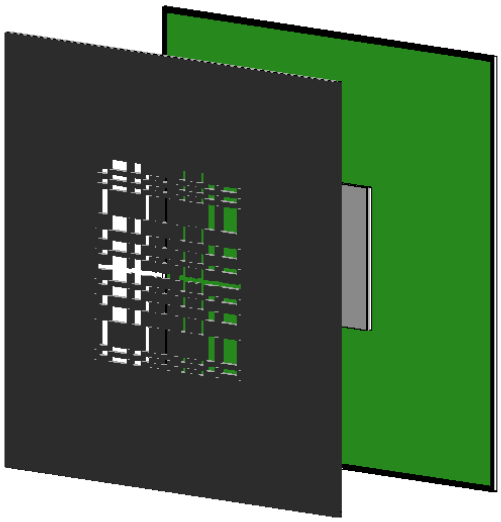


Figure 3.2: Mask placed in front of a position-sensitive detector

verification and optimisation will be needed to demonstrate this; Skinner [2004] concluded that coded mask telescopes

> "are indirect and inefficient... But nevertheless there are important regions of parameter space in which they are the preferred technique."

Since high-energy electrons and ions become increasingly difficult and resource-heavy to focus, energetic particle detection may be an area in which a mask would be advantageous in terms of resources and practical in higher-density regions or where the angular distribution of the particle radiation would not be expected to change significantly over a long enough integration time.

The design is further driven by the measurements which may be of use to space weather research and analysis on future swarm missions. Scientific fields that may be enlightened by measurements taken within the vicinity of a typical CubeSat orbit include those understanding the changes in particle populations at high time and

spatial resolution scales.

3.4.1 Design considerations

Within the family of mask and pixel detector solutions, whichever detector is available, a certain number of design points will need to be considered and finalised.

The decision on all of these parts of the design will be driven primarily by potential science return, balanced against efficiency, cost and signal to noise ratio, while bearing in mind the dangers associated with likely thermal, computational or telemetry issues, to produce an instrument capable of flying on as many different platforms as possible.

Detector choice

At a minimum, the detector should be position sensitive to the particles of interest, and able to avoid light and other sources of interference. In addition, the detector, along with the associated support systems (mechanical, electronic and thermal) will need to fit within the potentially very tight requirements of the spacecraft platform.

Geometry

The geometry design will largely be dependent on the distance between the mask and the detector, and on the shape of the mask pattern.

The computational simulations used to investigate the effects of these decisions will be given in detail in chapter 4.

Readout software and electronics

The readout electronics and operation depends heavily on the detector. It is to be decided whether individual particles should be counted or whether an integration time should be allowed within the detector.

While the time and pulse data of each event would allow further analysis, for example examination of data on shorter, or any selected, timescales, this would increase the data created significantly and in practice, it is likely that pre-selected integration times may be necessary to be feasible.

Depending on the detector and environment, it is possible that noise from thermal current and electronics might be significant. The operational behaviour of the system could be adapted to mitigate that, for example by reducing integration time to reduce thermal noise and identifying and summing incident particle positions.

Spacecraft requirements

Throughout the design process, the requirements on the satellite itself need to be evaluated. These requirements include consideration of the satellite's Command and Data Handling System (CDHS), telemetry, power, harnessing and any thermal control needed by the detector or other components.

3.4.2 Component selection

In this design, the component choices to be made that will have the largest effects on the behaviour in simulations, experiments, and performance will be:

- Detector type and geometry
- Mask shape, size and material
- Shielding considerations

Each of these will need to be decided, by lessons learned from previous missions or by theoretical calculations and simulations.

Choice of detector

The types of position-sensitive detector have already been discussed in sections 2.3.2 and 3.2.

The instrument geometry would be largely designed around the geometry of the detector. The mask size, distance and resolution are all closely related to the size and number of pixels of the detector.

The pixel size should be small enough to provide the highest resolution possible, however this is physically limited by the size of a signal produced by the impinging particles, which depends on their type and energy.

The fill factor, or active proportion of the detector area, should be as high as possible to maximise the number of particles detected and to reduce as far as possible the noise of the reconstructed angular distribution.

Ideally, the Medipix should be able to produce the most interesting data and is the most novel in this application. However, the analysis of the data is likely to be more involved if the opportunity of investigating individual particle hits is used to analyse particle types and fast changes in the particle environment. The simulations in chapter 4 are run with the geometry of a Medipix purely in an integration mode with an exposure time which will depend on the environment, the mask shape and the rest of the geometry.

Other types of CMOS detector would also be capable of providing good angular resolution and these have often been used in the past for charged particles. A CCD would also be an interesting choice: although these detectors are commonplace in a space environment, very few have been used specifically for the detection of electrons or charged particles, especially at higher energies. The use of a CCD is the method used in the proof of concept instrument tested and analysed in chapter 5, due to availability.

Choice of aperture shape and design

As described in section 3.5, there have been numerous studies on the optimum shape for a mask. In many cases, the choice is decided by balancing the need for the highest possible level of flux to the detector against the best shape for ideal analysis.

In theory, for X-rays, a Uniformly Redundant Array (URA) is the best choice for

the highest signal to noise ratio. The variety in the shapes and sizes (in grid design) into which URAs can be cut is discussed in section 3.5.2. This theory is not likely to be so relevant in a case where a thicker mask is required. It may be that this could be compensated for by a curved mask, by a mask shape with some distortion at the edges, or by the analysis algorithms alone. The simulations in chapter 4 will attempt to quantify these parameters.

The mask itself would ideally be as thin as possible, and so the material chosen for it should be as dense as possible - tungsten is likely to be one of the best choices.

Choice of foil material

If a foil were used for absorbing lower-energy particles, the energy range would be moved to higher energy particles and the count rate would be reduced in an LEO environment. Multiple materials have been used on previous missions: a beryllium window was used in REPTile; a titanium window was used on both of the solid state particle telescopes on Mariner 10; a nickel foil was used in the Medium Energy Proton and Electron Detector (MEPED) instrument on the National Oceanic and Atmospheric Administration (NOAA) satellites and tantalum and aluminium are used for shielding on Electron Loss and Field Investigator (ELFIN).

In the case of a position sensitive detector, this foil would be replaced with a coating on the detector or a thicker dead layer to avoid deflecting particles to other pixels to a greater degree than necessary.

Shielding considerations

While the low mass budgets on CubeSats and other small satellites reduce the amount of possible shielding, the geometrical dependencies in this design of instrument would suggest that the instrument would be best-positioned set back from the surface of the spacecraft. The side of the structure would be used as shielding and collimation to reduce the Field of View (FOV) to the limits of the coded FOV of the design.

Depending on the environment of high-energy particles, the possibility of bremsstrahlung created in the shielding may be a concern, and the layers of shielding added should aim to reduce this.

On the other hand, it should be borne in mind that sources which might be expected to illuminate the detector evenly will add noise and uncertainty to the measurements of the instrument, but not bias.

3.5 Coded apertures

The use of coded apertures in X-ray and gamma-ray sky astronomy was proposed and developed in stages throughout the 1960s. The technique gained popularity as a way of balancing geometric factor against higher angular resolution in low-flux images.

One of the earliest proposals of a mask-based technique for X-ray astronomy was in Mertz and Young [1961] who described a Fresnel mask, the shadow pattern of which could be optically deconvolved.

Such a concept was significantly developed by Dicke [1968] and Ables [1968]: a mask with a pattern of holes is placed in front of a position-sensitive detector, as shown in figure 3.3.

In the words of Caroli et al. [1987], the production of sky images in gamma and X-ray wavelengths in traditional camera systems was always hindered by

(a) the general difficulty in focusing high energy photons and

(b) the weakness of the fluxes with respect to the background counting rate.

This highlights the difficulty in the use of a pinhole camera method for certain situations: the smaller the pinhole, the higher the possible angular resolution per photon detected, but the lower the likelihood of photon detection at all. These random or coded aperture patterns are effectively increasing the possible detected

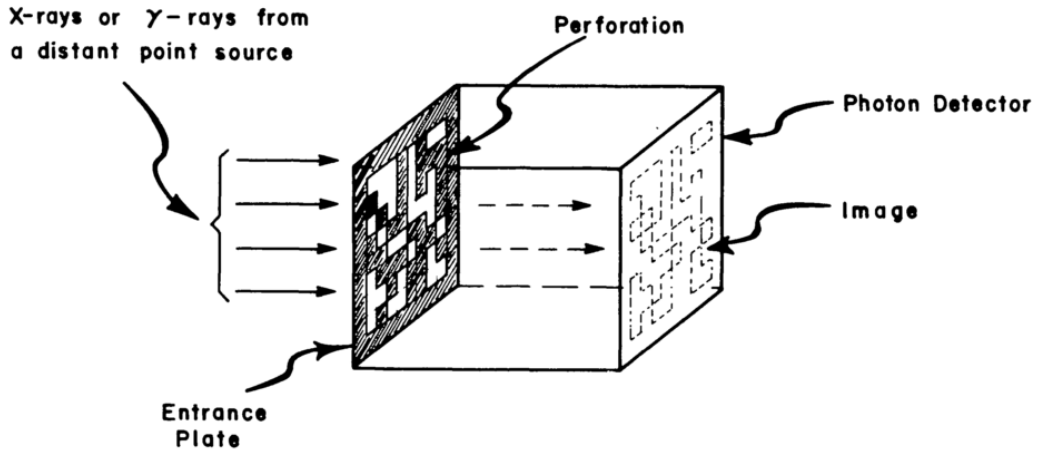


Figure 3.3: The original 'scatter-hole camera' concept for X-rays or gamma rays by Dicke [1968] showing a random pattern of holes in front of a detector.

signal: by using a position-sensitive detector and a mask with multiple pinholes or apertures, precise angular data can be retrieved while detecting up to half the local flux in the field of view.

The concept has been used and continues to be used in the field of X-ray and gamma-ray astronomy. Early applications included balloon or rocket-based instruments, for example those described by Gunson and Polychronopulos [1976], but they are used in current and future planned missions, in particular as X-ray and gamma-ray source monitoring, some examples of which are given in section 3.5.5.

Such a concept has also been used for several other, ground-based purposes, for X-ray and gamma-ray applications, but also for particle radiation. These technologies are highly relevant in the attempt to use the coded aperture technique for space plasma particle measurements.

For example, such a system has been used by Talebitaher et al. [2012] using a radioactive alpha source, to demonstrate its potential for use in imaging the products of plasma focus devices. Such a method could also be of use in diffraction experi-

ments, and is becoming more commonly used in single-photon emission computed tomography (SPECT) imaging systems, for example those used by Meikle et al. [2001].

There have been applications of the technique within neutron scatter facilities and the nuclear and security industry to locate and possibly identify radioactive sources from the detection of neutrons. For example, Vanier [2003] and Dioszegi et al. [2013] used a coded aperture with a gaseous neutron detector array to analyse thermal neutrons and Marleau et al. [2010] demonstrated a possible method for the location of fissile materials though high angular-resolution measurements of fast neutrons using liquid scintillator detectors.

Gal et al. [1997] and Gal et al. [2001] developed the first portable gamma source imaging spectrometer, based on a pinhole, scintillator and CCD, which is designed for uses in nuclear facility decommissioning. This design was then significantly improved by Gal et al. [2006], Gmar et al. [2011] and Lemaire et al. [2014] by the use of a coded mask in tungsten alloy and a Timepix chip. This allowed significant reductions in the exposure time required for a measurement, and a reduction in instrument mass from approximately 15 kg to 2 kg.

These more recent uses benefit heavily from the improvements in computing capabilities, which mean that there are now a range of feasible and effective deconvolution algorithms that can be used, and the electronics associated with the detector and the deconvolution are now small and low-power. The challenge in applying this concept to space plasma and radiation environments on a small scale is in part caused by the behaviour of electron interactions in detectors. If the environment contains a large proportion of high-energy electrons, the unpredictability of the size and position of the resulting signal from each particle means that data retrieval becomes more difficult. The aim of this research is to demonstrate the capabilities and limitations of this method in various space environments.

3.5.1 Coded aperture theory

The projected pattern created on the detector is given, in an ideal case, by the convolution

$$h(x,y) = \iint f(\alpha,\beta)\, m(x + d\tan\alpha, y + d\tan\beta)\, d\alpha d\beta \tag{3.1}$$

where α and β are the angles of the incoming particles in the directions of x and y, the positions at the detector, m is the pattern of the mask, f is the incident angular distribution and d is the distance separating the mask and detector. In a pixellated detector, this image is converted to an array representing the intensity at each point on the detector.

This represents a linearly independent shadow image of the mask for each incoming angle of radiation. Where incoming radiation comes from more than one angle or is distributed, these patterns overlap and the image on the detector must be deconvolved to analyse the contribution from each angle and recover the incident angular distribution. This process can be very simple or far more complicated, based on the available processing power, and how many inputs are taken into account, for example the assumed noise distribution or further assumptions about the behaviour of the source distribution or the detector. Some deconvolution techniques from the literature are discussed in section 3.5.3.

3.5.2 Coded aperture design

Initial instrument designs such as in Mertz and Young [1961] used a Fresnel mask for X-ray source identification. This created a pattern of a series of Fresnel zones on a photographic plate which could be scaled down and directly used to focus light onto a reconstructed image plane. However, with the advent of more advanced detectors and digital processing techniques, the nature of the masks used for these methods became based on gridded arrays.

Dicke's original paper focused on a randomly-perforated entrance plate, with

square holes manufactured on a regular grid. Randomly-designed apertures are a good starting point, since a key feature of an aperture design is the lack of repeating features which would cause ambiguities leading to image artifacts in the reconstruction of the data. However, other patterns are mathematically less sensitive to noise effects and will not produce artifacts as a result of their design, some of which are described in this section.

The choice of mask shape and scale is a balance between the need to increase the instrument's geometric factor, by increasing the open area, and the need to maintain the maximum ability to deconvolve the output to measure the environment with the best angular resolution. The optimal size, shape and scale of the mask will depend to a large degree on the requirements of the instrument - including FOV, angular resolution, mass and volume - and the size, shape and resolution of the detector. The expected environment will also be crucial for selecting exact design parameters - for example, the open fraction of the mask is important if the geometric factor needs to be maximised in low flux environments.

Another consideration for the design of physical masks, especially in space, is the mechanical rigidity of the structure, which is partly dependent on the material, partly on supporting structures and partly on mask shape and size. This structural durability can be achieved either by choosing a mask pattern with holes in positions which allow the material to be one continuous shape, for example the pseudo noise product arrays described by Byard [1992] or by adding further structural support between each element in the grid.

The sensitivity of coded mask telescopes can be complicated to quantify analytically. Often assumptions are made whereby the supporting structures and defects in the grid and detector pixel array are ignored and noise is subject to assumptions of a Gaussian distribution. These assumptions should not be ignored - simulations can replace analytical calculations, or the calculated value can be modified by the improved methods described by Skinner [2008] in his critique of the standard numerical estimates used for sensitivity.

Uniformly Redundant Arrays

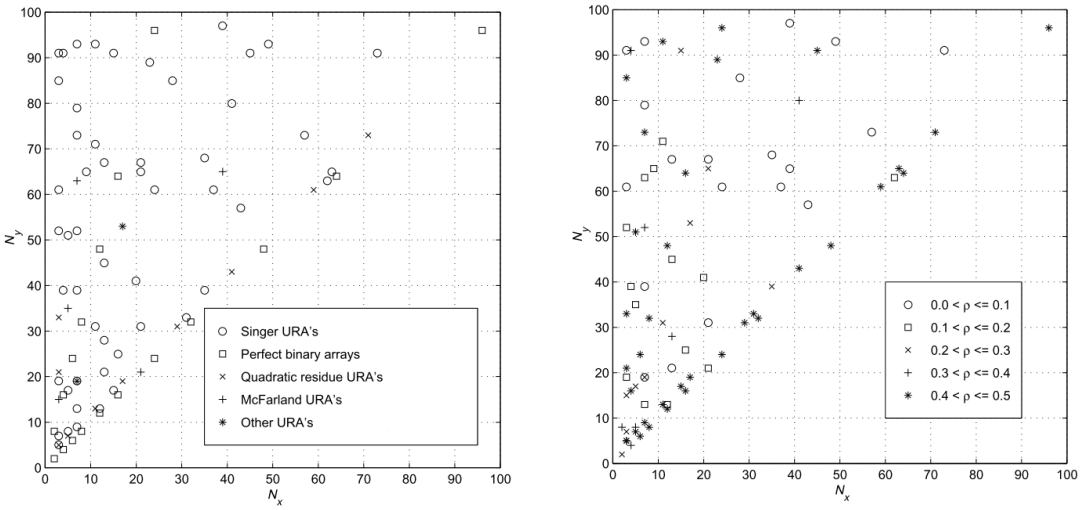
A mathematically ideal mask pattern is an array which has an autocorrelation function with perfectly flat sidelobes; such a pattern is known as a URA. With these patterns, a deconvolution of the retrieved frame with the mask shape will have a minimised level of noise and therefore URAs are capable of giving far better SNRs than random arrays. Some simulated examples of the huge improvements in performance and capabilities by the use of URAs are described and demonstrated in Fenimore and Cannon [1978] and Fenimore [1978].

These patterns can be constructed in one dimension or two dimensions, for use in measurements in one or two directions and depending on the geometry of the detector. One-dimensional patterns have been constructed for X-ray detectors, for example the cyclic difference sets described in Gunson and Polychronopulos [1976]. 3.5.5.

For two-dimensional aperture patterns based on square pixels, Busboom et al. [1998] have summarised the known URAs and patterns related to URAs up to dimensions of 100×100 pixels, which are shown in figure 3.4 along with the open fraction of these patterns. A brief summary of the types of patterns available is given here.

Examining the quadratic residues of finite fields can be used to produce quadratic residue URAs, as described by Fenimore and Cannon [1978] and Calabro and Wolf [1968]. Fairly simple cases give almost square arrays of a range of sizes, while other choices can lead to strongly off-square URAs; the open fraction is typically near 50%.

There is an equivalence between URAs and abelian difference sets so families of difference sets are useful to consider. One such important family is the Singer difference sets first described by Singer [1938], arising from cyclic groups based on projective geometries in finite fields of prime or prime power order. The associated Singer URAs give a wide range of choices, some of which are almost square arrays with an open fraction of about 50%. Other versions with lower open fractions can



(a) URAs by construction method

(b) URAs by open fraction

Figure 3.4: Known URAs up to 100×100 pixels from Busboom et al. [1998]. N_x and N_y are the sizes of the array in each direction.

be useful when the detector resolution is in the order of magnitude of the mask element size. These were described by MacWilliams and Sloane [1976].

McFarland [1973] described a family of URAs arising from non-cyclic, abelian groups with some cases leading to square URAs which, in general, have low open fractions and therefore have limited applications in coded aperture imaging.

The Perfect Binary Array (PBA) family is a family of square or 1:4 aspect ratio URAs with a 50% open fraction, where the sides are powers of two. These arrays were developed by Lüke [1987], Jedwab and Mitchell [1988] and Bömer and Antweiler [1990]. The iterative method of building them is described in Busboom et al. [1998]. A PBA was used in the experiments described in chapter 5 because of its large open fraction, square shape and grid size that is a factor of the pixel number of several detectors.

A family of arrays related to URAs is described by Gottesman and Fenimore [1989] - the group of modified uniformly redundant arrays (MURAs). Although the aperture array and its inverse filter are not identical, the difference is small, particularly for large arrays, and their square size and open fractions of about 50% make them a good choice when URAs are not available for a required size.

The choice of type and scale of URA or similar mask patterns needs to be made based on the planned detector size and resolution, the required open fraction and the geometry of the instrument.

Hexagonal arrays

URAs and MURAs can also be fitted to a hexagonal array such as the family of skew-Hadamard Hexagonal URAs (HURAs) described by Finger and Prince [1985], an example of which is shown in figure 3.5 - these shapes benefit from greater symmetry and could be an improvement on a square array in the case where the detector is position-sensitive in a circular or hexagonal pattern, for example, the proposed hexagon-fitted Medipix described by Llopart et al. [2001].

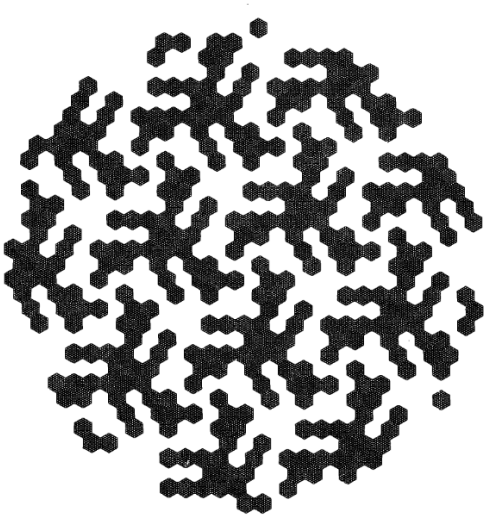


Figure 3.5: An example of a hexagonal URA of order 67 from Finger and Prince [1985] where white represents an open cell and black represents a closed cell.

Other mask designs

Each of these designs could also be scaled, in one or both directions, to produce a different FOV and resolution in each direction. In addition, the holes themselves could be any shape or size, for example to simplify manufacturing, to change the shape of the resulting Point Spread Function (PSF) or to increase or decrease the detected flux.

More generally, a reference structure need not be two-dimensional. The mask is not restricted to being flat, although the post-processing of the resulting data would rapidly become more difficult with the extra complexity. Potuluri et al. [2003] have described a general mapping of radiation position and angle to pixel position, using an example of a pattern of light travelling along angled light pipes.

An attempt to design holes at various angles to allow flux from a wider range of angles was made by Hong et al. [2004]. This could be of interest in the case

of the detection of high-energy particles where a thicker mask is needed to fully absorb the particles' energy. However, such a development would make the design, simulation, manufacturing and deconvolution significantly more complex, since the mathematical benefits of a URA are no longer straightforward in this case, so for the lower ranges of energetic particles, a flat mask is preferable.

3.5.3 Deconvolution methods

Deconvolution is the process of decoupling a source image from its PSF. Such a technique has a very large number of uses in astronomy where in almost every case a system will form an image of detected photons or other particles which is only a representation of the useful data.

The output of a coded aperture instrument based on a digital detector will be an array of digital numbers which must be processed in order to retrieve the angular spread of incoming particles. In this case, the shape of the mask itself is a large contributor to the PSF. This can be scaled for viewing point sources closer than infinity, but for relevant space plasma applications, the particle flux at each angle is of more interest, and to provide this the mask pattern should not be scaled.

Cross-correlation

A simple deconvolution algorithm would estimate the value for each pixel in the calculated image by cross-correlation - moving the mask over the data image and summing the product of the mask and image for each pixel.

For the ideal case given in equation 3.1 the reconstructed object is found by the cross-correlation

$$F(\alpha, \beta) = \sum_i \sum_j H(i,j) M(i+\alpha, j+\beta) \tag{3.2}$$

where H is the image in the detector plane, and M is the mask shape at the same

scale, and both are in the form of an array. This produces an array of flux intensity which can be mapped to the incoming angles. The scaling of the numbers in the array to flux can be achieved by scaling the values used to represent open and closed sections of the mask, and by scaling the final flux map. The exact scaling will depend on the detector type, the particle energies and the exact setup. In practice, this scaling must be determined by experiment, since the signal produced by each particle reaching the detector can only be known accurately from empirical measurements. This method will give a map which includes the PSF of the shape of the holes in the mask.

Improved deconvolution algorithms

More advanced algorithms can produce more accurate results by taking into account the expected distribution of noise in the data and reducing the number of necessary assumptions. In addition, there is often a need to reduce the number of computations required when on-board processing is both required and restricted and the frame rate or frame size is high. Compression of frames before downlinking may also be necessary, and algorithms that require less telemetry are also of interest.

General deconvolution including a noise component can be described in a simple form, after Fenimore and Cannon [1978] and Braga et al. [1991]. Including detector noise, the form of the image on the detector is

$$H = F * M + D \tag{3.3}$$

where $*$ is the correlation operator, H is the image, M is the aperture shape, F is the angular flux distribution and D is the detector background noise, all in array form.

If the mask shape M is exchanged for a post-processing operator,

$$\begin{aligned} \hat{F} &= H * O \\ &= F * (M * O) + D * O. \end{aligned} \tag{3.4}$$

The operator O can be chosen to optimise the $\hat{F}$ reconstruction.

If the operator O is chosen so that $M * O$ is a delta function,

$$\hat{F} = F + D * O \tag{3.5}$$

where $D * O$ is the noise term in the reconstructed image. This is the case where a URA is used as the mask and $M = O$. However, the operator O and processing can be adapted to reduce this noise component.

Fenimore and Cannon [1981] described reconstruction techniques using Fourier transform-based methods, which are implemented using fast Fourier transforms (FFTs) with a range of sampling patterns. Although it was demonstrated that large FFT arrays were required relative to the size of the image frame, the computational power required is quickly reduced relative to cross-correlation, the larger the image.

The Lucy-Richardson algorithm from Richardson [1972] and Lucy [1974] is a Bayesian-based iterative method of deconvolution developed with the assumption that the noise is Poisson-distributed. This is capable of significantly reducing the smearing associated with the PSF intrinsic to the individual holes in the mask, and converges after multiple iterations. The modified version designed for coded apertures discussed by Sambo et al. [2009] bases each estimate F^k on the $(k-1)^{th}$ estimate from

$$F^k = F^{k-1} M^T \frac{H}{M \cdot F^k + D} \tag{3.6}$$

where the notation is defined as above.

A series of fast decoding algorithms was developed by Roques [1987] for use with URAs, which compute the standard cross-correlation but with fewer calculations by removing the need for multiplication steps which are known to be zero. However, this is unlikely to be necessary with modern hardware for the applications used here. For applications such as medical imaging or radioactive contamination mapping where speed is key, these techniques are still important, and Byard [2014]

has produced similar algorithms for MURAs which can reduce the processing time by approximately 25%.

3.5.4 Coded aperture techniques

Techniques related to coded aperture imaging use the coded aperture concept, but add an additional step to increase the capabilities compared to a basic mask. However, this is very likely to add to the complexity, mass or processing required for the instrument, and would be difficult to integrate into a small-scale instrument for a small satellite platform; these techniques have been noted and considered for further development compared to the initial designs described in the rest of this chapter and chapter 5.

Several instruments have used the rotation of the spacecraft or of a rotating modulation collimator to analyse the signal in the detector or detectors over time. The All-Sky Monitor (ASM) on the RXTE mission, described by Levine et al. [1996], used independently-moving cameras with one-dimensional masks at 90° from each other to scan a large field of view for X-ray sources.

Methods for improving the signal to noise ratio further for weak sources were tested using a mask-antimask setup, where both a mask and its inverse were used, which allows differences in behaviour across the detector to be filtered out. One example is the balloon-borne Gamma-Ray Imaging Spectrometer (GRIS) described by Gehrels et al. [1985] and Teegarden et al. [1985]. This used motors to rotate four separate mask components to create a full antimask, and could also offset the mask. Braga et al. [1991] described Telescópio IMAgeador de raios-X (TIMAX), another simple coded-mask technique used on a balloon-mounted X-ray telescope where the pattern of the mask is effectively inverted. This was achieved for a full grid excluding the central element by rotating by 90° a quadratic residue URA-based coded mask. This was demonstrated by Braga et al. [1994] to be very successful in reducing systematic variations in signal during a calibration test on a balloon flight.

Several types of URA fit or closely fit the symmetry requirements for these

methods, and Byard [1992] described the construction method of two families of 90° antisymmetric MURAs for applications such as these. Busboom et al. [1997] described the possible improvements in the general reconstruction of coded aperture data with measurements using more than one arbitrary mask shape.

This technique was also described and demonstrated for a hexagonal grid shape by Cook et al. [1983] and Althouse et al. [1987] using a hexagonal URA which is close to being antisymmetric under 60° rotation. Each of these techniques provide significant SNR improvements for their setups, but the gains involved depend largely on the behaviour of the detectors used. Such a concept could however be feasibly fitted within a small instrument package.

An interesting possibility for wide-band X-ray and gamma-ray imaging was proposed and simulated by Skinner and Grindlay [1993]. A thick coarse grid capable of stopping all incident radiation contains a thinner array within each hole in the mask, which is transparent to lower-energy photons only. The use of detectors capable of identifying the position and energy of the incident radiation means that high-resolution coded measurements can be taken of higher energy photons and higher-resolution coded measurements can be taken of lower energy photons. However, adapting this concept for a particle instrument would present significant challenges owing to the deflection of light particles when they pass through thin materials.

3.5.5 Space telescopes making use of coded apertures

The use of a coded aperture is often, as discussed in the previous section, prompted by the need to find high angular resolution information about a source, with radiation of a type it is hard to focus, with as little loss of flux as possible. This approach has therefore been in use for hard X-ray and soft gamma-ray measurements, especially those searching for distant point sources, such as Gamma Ray Bursts (GRBs), over a very wide field of view. The two examples given here illustrate the types scales, materials and detectors which are typical in these current

space applications.

INTErnational Gamma-Ray Astrophysics Laboratory (INTEGRAL) imagers

The Imager on Board of the INTEGRAL Satellite (IBIS) and Spectrometer on INTEGRAL (SPI) instruments were launched in 2002 on the the INTEGRAL satellite and are designed to observe hard X-ray and soft gamma-ray targets between approximately 15 keV and 10 MeV photon energies. In IBIS, a URA coded mask and two pixelated detectors in a telescope arrangement are used - this setup is illustrated in figure 3.6. It has a fully-coded FOV of $8.3° \times 8.0°$ and an angular resolution of 12 arcmin full-width half max (FWHM) (Goldwurm et al. [2003]). As is typical for space-based high-precision X-ray telescopes, the instrument is on a large scale and has a mass of 677 kg.

The Wide Field Monitor (WFM) on Large Obervatory For X-ray Timing (LOFT)

LOFT is a proposed ESA mission, which could be launched around 2025. One of the two instruments it would contain is the WFM described by Brandt et al. [2014] and shown in figure 3.7 - a coded-mask wide field X-ray monitor with five cameras, each of which is made of a pair of silicon drift detectors coupled with masks. Since a linear drift detector has a high resolution in one direction only, the mask is correspondingly designed for the same direction, with a 25% open fraction, and the second camera in the pair works in the perpendicular direction so that a point source direction can be pinpointed to an accuracy of $1'$ and the general angular resolution is $5'$. Although this instrument is designed to be on a smaller scale than IBIS, above, it still has a total mass of 125 kg, several orders of magnitude larger than the concept to be developed here.

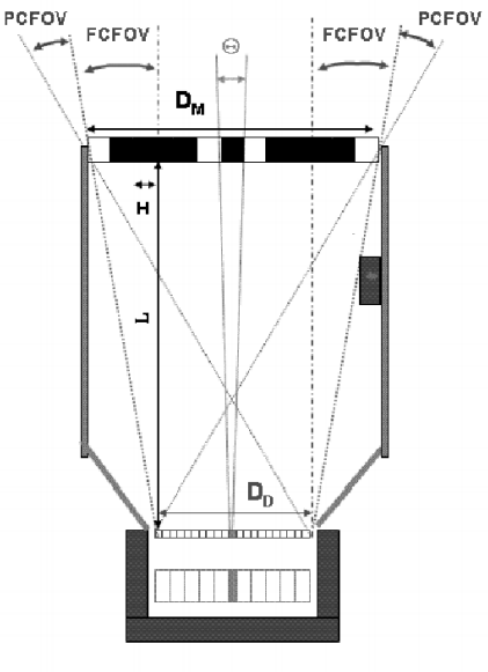


Figure 3.6: Schematic of the IBIS instrument on INTEGRAL from Goldwurm et al. [2003] showing the limits of the fully-coded field of view (FCFOV) and the partially-coded field of view (PCFOV).

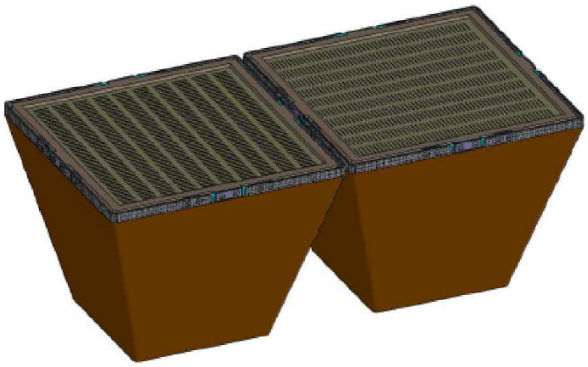


Figure 3.7: Computer-Aided Design (CAD) of one of the ten detectors in WFM proposed for LOFT, as shown in Donnarumma et al. [2012]

Chapter 4

Simulation setup and results

Before the proposed geometry and parameters for the coded aperture instrument were chosen, the concept was reviewed for mathematical and historical background and then modelled with different geometries using the Geant4 framework described in section 4.1.1. The most important aspects of the design are then integrated into the lab-based experiments described in chapter 5 and the proposed flight instrument in chapter 6 and the possibilities for further development in chapter 7.

4.1 Simulation environment

As was mentioned at the beginning of section 3.5, the development of a space particle detector based on the coded aperture concept will depend largely on the ability to filter and accurately detect the position of particles. The first stage of this development is the computational simulation of particle propagation, transport and energy deposition through a series of models.

Such simulations of particle transport and interaction are commonly used in space instrumentation, both for detector development and shielding analysis, among other applications, but they are also used in ground-based particle detectors and accelerators. Throughout the history of all such experiments and developments,

predictions about the behaviour of such instruments have been needed in advance of their production. The capabilities of computers have quickly overtaken analytical calculations in convenience and accuracy, and the simulation environments available to researchers have increased in sophistication, speed and accuracy over time. The toolkit used in these simulations - Geant4 - is a powerful and often-used set of software for space radiation simulations.

4.1.1 Geant4

The first version of the GEometry ANd Tracking simulation toolkit (Geant) was developed at CERN in 1974. Geant4 was a significant development in 1993 from the previous iteration, Fortran-based Geant3; the Geant4 simulation toolkit and related applications are used in almost all simulations of the passage of particles through matter in space radiation applications. Therefore, there is a huge user-base of application and tool-developers for almost every environment and situation.

Other applications based on Geant4 have been developed by others - in particular Mulassis, which was first created at QinetiQ in 2002 (Lei et al. [2002]), updated in line with Geant4 releases until 2010 and was available to download until 2013. Mulassis is designed for shielding analysis for spacecraft; a number of layers can be set up and the particles to simulate are defined. This can also be of use in simple geometries where the output needed is simple energy attenuation and deposition. Mulassis, and other applications, have been made available to be used on a remote server at www.spenvis.oma.be, as was used in section 3.3.1.

Both Geant4 and Mulassis have been used in the following simulations to some degree - Mulassis to quickly obtain a broad overview, for example measuring the energy deposition distribution in silicon, and Geant4 to extract more complete data from a more complicated and accurate geometry in three dimensions.

The simulations here have been run using Geant4.9.6. Geant4.10.0 was released in 2013, with improved parallelisation capabilities and some updated libraries. For the relatively simple simulations involving electrons and protons described here, the

main expected result of these updates is a small-scale increase in speed and very little difference to the operation or results can be expected; therefore the results from the simulations in Geant4.9.6 were used.

Monte Carlo particle simulations

A Monte Carlo particle simulation generates a large number of particles distributed pseudo-randomly across the parameter space, for example position and angle, and performs a deterministic calculation of their behaviour. This method is highly effective for many-dimensional parameter space, where the method quickly converges and produces a lower (and calculable) uncertainty and a lower risk of artifacts in the data compared to an attempt to evenly sample the parameter space.

Geometry definition

Geant4 applications are written in C++. The geometry is defined as a series of volumes, which must not overlap, and the particle starting positions, angles and energies are defined.

The description of a series of simple volumes, such as cuboids, cylinders or spheres with dimensions, positions, rotations and materials used in a Geant4 application is specified in Geometry Description Markup Language (GDML) format, which is based on Extensible Markup Language (XML).

Although it is possible to export a GDML geometry from a CAD format or program, this becomes much more complicated and time consuming for more detailed models, as well as significantly slowing the simulation process down, so it becomes necessary to choose only the most important components of an experiment for a particular simulated process, based on what will have the largest effect on the sensitivity and shielding.

The geometry is defined within C++ using the GDML principles. Each volume's shape is defined individually and then it is placed with a defined translation and rotation relative to its parent volume; this creates a hierarchical structure of volumes.

A material from the Geant4 material library is used to describe the volume's properties. The definition takes place in a class inherited from the detector construction class, which is called to generate the user geometry and to return a handle to the world volume which contains all other geometry. A scoring volume is selected in which any energy deposition is logged for post-processing - in these simulations this is the active volume of the detector.

The source is defined in a class which specifies the particle type and kinetic energy as well as the starting position and particle direction relative to the coordinate system of the world volume.

This geometry and source definition is part of a full application which includes the required physics libraries, visualisation, and simulation management.

Tracking algorithm

Geant4 uses Runge-Kutta methods to form a close approximation to the path a particle takes through a field, and transports the particle through a material based on interaction probabilities.

The statistical distribution of interaction lengths before a new energy and trajectory is chosen is calculated based on the probabilities of predicted interactions within the material. Using the notation from the Geant4 physics manual Wright [2011], the mean free path, λ, of a particle with energy E can be given in terms of the total cross section:

$$\lambda\left(E\right) = \left(\sum_{i} \left[n_i \cdot \sigma\left(Z_i, E\right)\right]\right)^{-1} \tag{4.1}$$

where $\sigma\left(Z, E\right)$ is the total cross section per atom of the process and $\sum_i$ runs over all elements composing the material.

An interaction point is determined using a statistical distribution of the mean free path of the process. The number of mean free paths a particle travels is independent

of the material through which it is travelling and is given by

$$n_\lambda = \int_{x_1}^{x_2} \frac{dx}{\lambda\,(x)}, \tag{4.2}$$

which is used to calculate the distance from one interaction to the next.

In addition to the given distance between interactions, different physical processes can impose limitations on the maximum step length if this is required to prevent loss of accuracy of the simulation.

Geant4 operates in steps to simulate the behaviour of a particle. The distance between the steps for a given particle is described above and is used as long as the particle does not enter a different volume within this distance. It differentiates between continuous, discrete and resting processes, or a combination of these. A continuous process will affect the particle in parallel with the other continuous processes, but only take cumulative effect on the trajectory and properties of the particle after all continuous processes are complete. The discrete process will take effect immediately after its invocation. A resting process also takes effect on a resting particle, with nuclear decay being an example for a combined discrete and resting process.

All processes with a continuous part such as the physical transportation or energy loss due to ionisation are executed after the determination of the step length; the change in kinetic energy only takes effect after the completion of all continuous processes. Position and time will then also be updated.

Should the particle have not been terminated by the continuous processes, then processes with a discrete part such as Compton scattering are invoked and immediately affect the kinetic energy and trajectory of the particle. Should any secondary particles get created (for example due to bremsstrahlung) these would be created at this point, and it is checked whether the track terminates. To complete the step, a callback function defined by the user is invoked, for example to save the hit information, and the trajectory is saved.

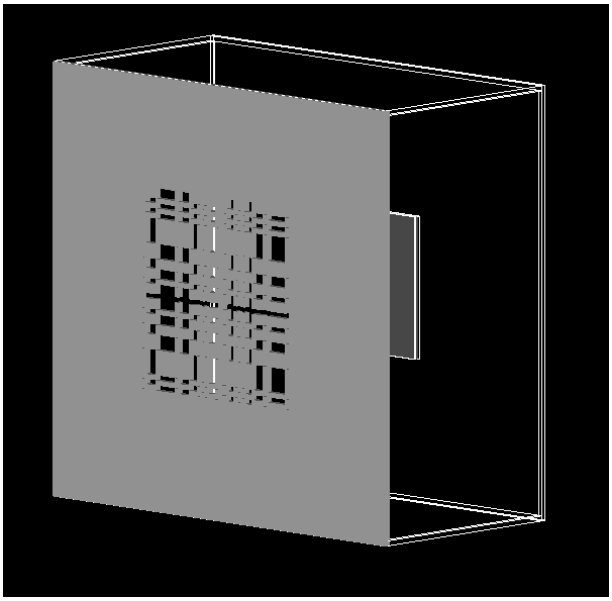


Figure 4.1: An example of a simulation setup in Geant4, showing the mask with a coded aperture pattern in the foreground with the detector behind and transparent shielding surrounding the instrument.

4.2 Mask simulations

These simulations aimed to optimise the parameters of the mask setup and prepare for experimental testing by providing numerical results for comparison with the further aim of identifying any as-yet unpredicted effects.

4.2.1 Setting up the simulations

A separate application was built for each simulated situation.

An example of an instrument setup in Geant4 is shown in figure 4.1. In this case, the mask plate is 4 cm×4 cm, while the side of the aperture is around 2 cm long. The mask is positioned 2 cm in front of the detector which is a square with an area of approximately 2 cm^2.

To set up the mask, the URA grid was calculated in Matlab and then exported via a binary file to the Geant4 application. Each position on the mask sheet is either represented by a cuboid of tungsten or an empty space. There is shielding at the sides and behind the detector. The detector itself consists of three blocks of silicon: the entrance window, the active region and the exit window. The block representing the active region is marked as a scoring region; therefore, all interactions taking place in that region are recorded during the simulation. Data from this region can be extracted; in the case of these simulations, the position (in three dimensions) and energy deposited at each interaction was recorded. This information was added to a comma-separated variables (CSV) file with no distinction between impinging particles - just the energy and position of each interaction is recorded, since for image analysis and deconvolution the energy per pixel over a chosen integration time are the data of interest.

Since the estimation of the distribution of the angles of the incident particles is the aim, for the simulations, the test source is chosen by angle in two directions and only particles that are aimed in the direction of the detector are flown. The angular distribution can be chosen according to what is most useful for the simulations - for example, isotropic, point source or linear. For the purpose of the demonstration simulations explained here, simple angular distributions were flown, with and without noise.

Figure 4.2 shows the behaviour of particles propagating from the left of the image, including one secondary photon (in green) created by an interaction between an energetic proton and the tungsten mask.

Once the list of energies and positions had been collected, further analysis was done in Matlab. For each interaction datum, the energy involved was assigned to one of the 256×256 pixels to form the overall energy deposited per pixel in the simulated integration time.

This process is shown in figure 4.3 and figure 4.4 for an example of irradiation by 300 keV electrons in a Geant4 simulation. Each particle reaches the front of the

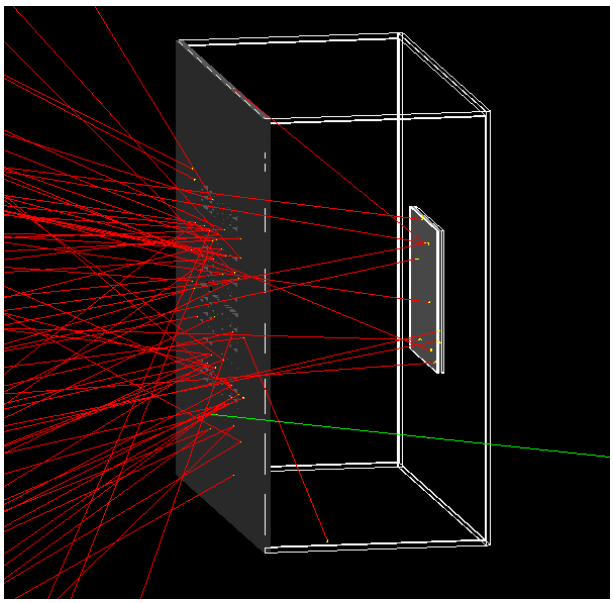


Figure 4.2: An example of an irradiated simulation setup in Geant4, showing the mask with holes on the left and the detector on the right, with the red lines showing the paths of electrons. The green line represents a photon created by bremsstrahlung within the mask.

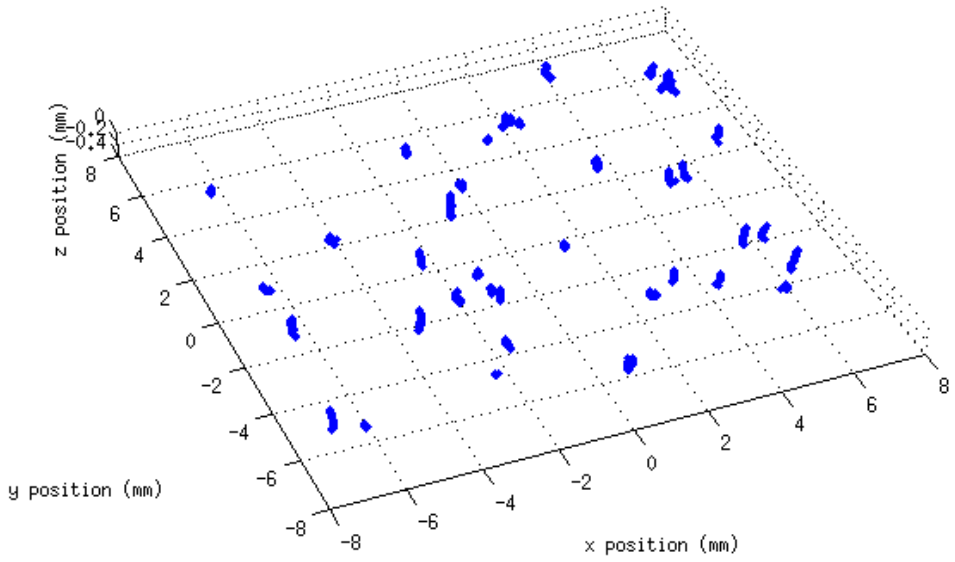


Figure 4.3: Position of electron interactions in the active region of the detector for 300 keV electrons

active region of the detector and passes into or through it.

Figure 4.3 shows the points of interaction calculated by the simulation. These are where an incident particle has interacted with the silicon lattice and a new direction has been calculated. Energy is deposited in this position as a result and this is recorded, but the plot shows the position only, not the magnitude of the energy deposition. Figure 4.4 shows the summed energy deposited in each pixel of the detector as a result of these interactions, demonstrating the transformation of track data to a two-dimensional output data frame.

These simulations have not taken into account the effect of charge sharing. When a particle deposits energy, especially further away from the collecting electrode, a

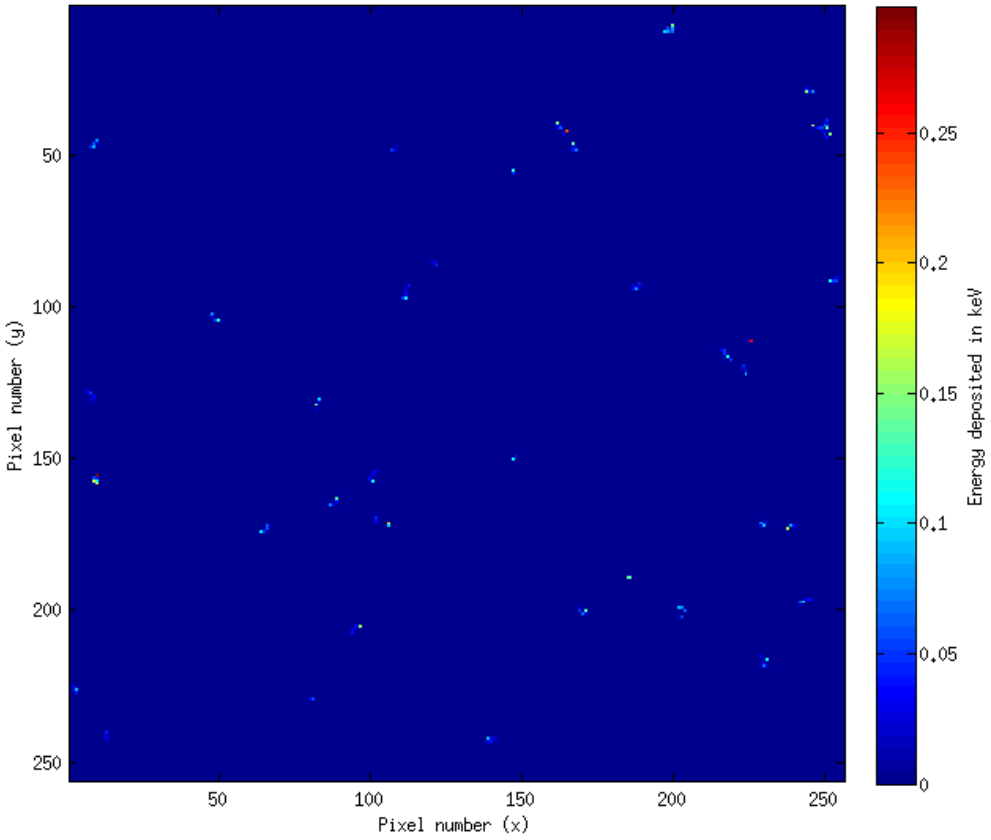


Figure 4.4: Summed energy deposition per pixel for the particle paths shown in figure 4.3

cloud of electron-hole pairs is formed and several pixels may record the interaction. Sosin [2012] determined analytically a model of an expanding spherical Gaussian cloud and suggested parameters that could be used in his model, including mobility and diffusion coefficients. Such a model could relatively easily be added to the analysis, since the interaction has been recorded in three dimensions and the distance from the event to the electrodes can be calculated, but the size of the effect will depend heavily on the type and geometry of detector. The simulations in this section used a detector with a Medipix-like geometry and mid-energy charged particles, in which this effect is limited to less than one pixel spread.

4.2.2 Simulations of the different shapes and sizes of mask

Many variations on the theme of the mask have been studied in the past. Initially, at least, the study here is confined to static two-dimensional mask designs of a constant thickness, since the design aims to be simple and low-resource. In this case, the variables to be considered are the shape, size, material and thickness of the mask.

Mask shape

As discussed in section 3.5.2, there has been extensive consideration of the best possible shapes of masks for previous instruments. Although the properties of these shapes remain mathematically optimal in the case of a small detector, the change in geometry needed for a miniaturised particle instrument should be taken into account. The mask, and therefore the holes in the mask, must be small, while the mask material must remain thick enough to stop particles that hit it, so the holes will be relatively long and narrow. Therefore a degree of collimation will occur and the detected flux from wider angles will be decreased. This effect is unwanted and must be balanced against the need to miniaturise, or avoided by adding additional collimation around the instrument.

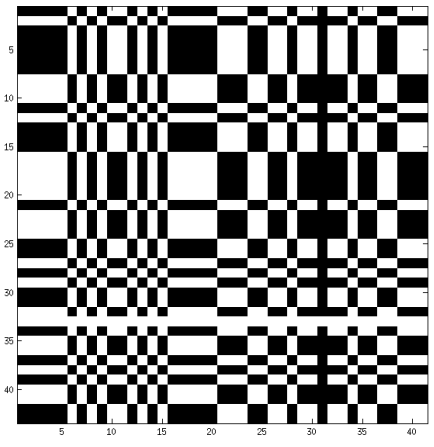
The mask shapes shown in figure 4.5 have been simulated. Since the possible detector geometries, such as a Medipix, were exactly or nearly square, masks which were close to being square, with a high open fraction, were considered. These included the quadratic residue URA family and the PBA family. The PBAs in particular are well matched to the Medipix detector owing to the grid size of either 16 or 32 being a factor of the 256-pixel size of the detector.

The types of mask grid that have been considered are shown in figure 4.6 using the example of a 41×43 URA mask shape.

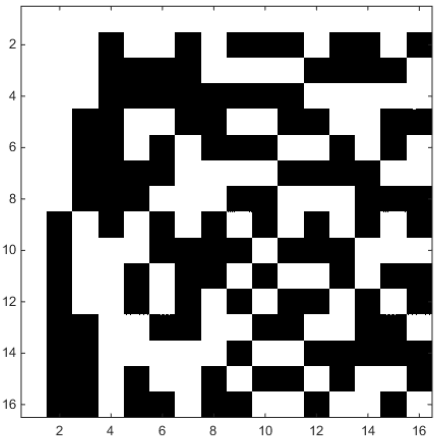
Shape (a) shows one of the URA patterns modified to be physically practical, in this case by moving the sides of the cut-outs inward, but similar methods could be used, for example thickening the join between two sharp corners of mask material or simply cutting a square for each open hole (leaving metal lines between each one). Shape (b) has taken this further: in this case, each open grid element is represented by a hole at half the scale of the original element. For this mask, this results in a complete open fraction of 0.125. This reduces the flux allowed through, and therefore the geometric factor, but improves the rigidity of the mask, and prevents incident particles at an angle travelling through effectively two holes, which is a situation which is not accounted for in the processing. Whichever method was used, the simulation of a grid of type (b) was modified by replacing each open element with three closed elements and allowing the final segment to be open; this caused the simulation setup and run to be slightly more time consuming than a grid of type (a).

Shapes (c) and (d) show the repeated mask shapes used by many including, for example, Fenimore and Cannon [1978], which increases the FOV of the instrument by approximately a factor of two in each direction. These are shown with an open fraction of both 0.5 (in (c)) and 0.125 (in (d)).

(a) Random



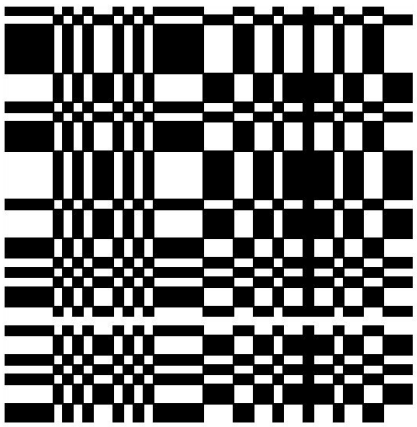
(b) 41×43 Quadratic residue URA



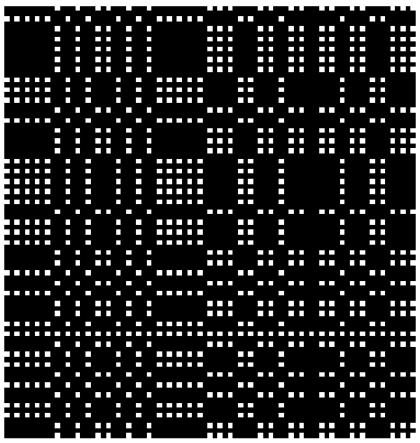
(c) 16×16 PBA

(d) 32×32 PBA

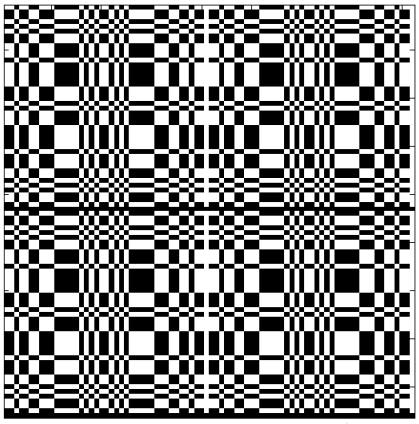
Figure 4.5: The array patterns for the mask that have been compared in Geant4 simulations



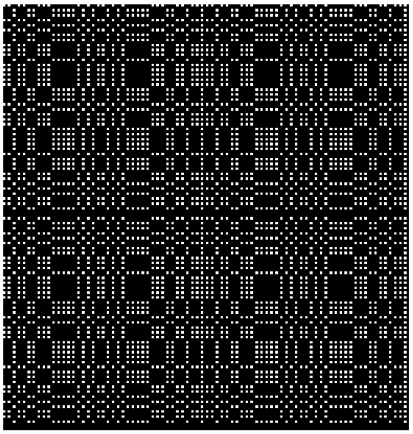
(a) Practical URA



(b) Smaller holes



(c) Repeated URA



(d) Repeated with pinholes

Figure 4.6: The mask construction types that have been compared in Geant4 simulations, shown based on a 41×43 URAs.

Mask size

The sizes to be chosen in the mask include the size of each hole and the number of squares (or other shapes) on the grid. The mask size should be the same as the detector area, or twice the size in each direction in the case of a repeated mask.

Optimised mask shapes, such as URAs, are limited in the grid size they can fit on. A smaller-scale grid will allow a greater angular resolution, but will quickly run into geometrical problems when the grid is on the same scale or smaller than the thickness of the mask since the incident particles are collimated to some degree. The simulations have been run using mask shapes with pixel sizes of between 300 µm to 900 µm and varying mask pattern sizes - the random mask was on a grid of 40×40 pixels, the URA was 43×41 pixels and the PBA was 16×16 pixels. In addition, these mask patterns can be repeated in each direction to give four times the number of holes at the same scale. As the mask is made larger, the field of view is improved, but the distortion of the PSF at greater angles becomes a greater problem. This could be improved to some degree by moving the plane of the mask further from the plane of the detector, increasing the angular resolution while reducing the field of view.

Mask thickness

For these simulations, the thickness of the mask has been taken to be 0.5 mm of tungsten. The STopping power And Range tables for electrons (ESTAR) and STopping power And Range tables for protons (PSTAR) estimate this thickness of tungsten to eliminate electrons above around 2 MeV and protons above around 20 MeV. Therefore this thickness and material for a mask should be appropriate in environments where there are only very low fluxes of particles above these energies.

In some environments where there are very energetic particles, the likelihood of bremsstrahlung caused by the tungsten mask should be considered, since the situation might be improved by a slightly less-dense material or a less dense layer on the front of the mask. These simulations take into account bremsstrahlung for the

simulated particles, and assume a low number of high-energy or cosmic ray events. When these high energy events do occur, they would be assumed to be at a random position on the detector and add to the background noise rather than cause a bias in the deconvolved data.

Deconvolution

The images directly acquired do not immediately represent the original angular data; deconvolution is required. Some of the deconvolution techniques employed by previous authors have been described in section 3.5. Whichever method of deconvolution is used, the mask image must be scaled to the physical size used in the simulation. Once the image has been deconvolved, the reconstructed original pattern should be recovered, although it can be expected to be noisy. Although this is not always the case, strong features can be easily identified with simple deconvolution algorithms. With more work on improving design and data analysis, these results are significantly improved.

The results of these simulations were processed largely with a straightforward cross-correlation and with the Matlab implementation of the Lucy-Richardson algorithm. The number of iterations of the Lucy-Richardson algorithm for the result to visually appear to have converged was approximately ten.

4.2.3 Simulating the source

During the full development of an instrument, both the test sources and flight environments need to be simulated. Some relatively simple source geometries have been simulated in this chapter to test and demonstrate the capabilities of the proposed geometry, for example to find the separation angle at which two near-parallel particle fluxes can be resolved and the capabilities of the instrument to analyse broader features.

The laboratory-based experiments described in chapter 5 also required some simulations to understand their behaviour, which are detailed in section 5.6. The

sources used for these simulations are based on mechanical knowledge of the sources, and physical predictions of the radiation energies and distribution.

For a radiation belt instrument, the simulation of the flight environment would include the predicted environments of the SAA and auroral regions. These environments have been simulated and described in chapter 6 with the aim of making the best estimate of the radiation environment so that comparisons between simulation and flight results would be as realistic as possible.

In most simulation cases, the particles to fly were selected based on their relevance. Although the particles were chosen from a full random distribution of whichever environment was needed, only those which were aimed at or around the detector were flown in order to increase the simulation efficiency.

4.3 Simulation results

A large number of simulations were run with different geometries and source distributions, to inform the plans for an instrument based on coded mask designs. As a starting point, the simulations were run using the case of an instrument based around a geometrically Medipix-like detector. This would then be adaptable for instruments that would use a different position-sensitive detector.

These simulations covered a range of different source distributions, with the aim of assessing the instrument capabilities and optimising the design. Some of the situations that were considered to help understand the instrument included measuring separation of sources, distinguishing sources in a background, and measuring distributed sources.

With a few exceptions, all the simulations were run using electrons as the source particles. The reasons for this were both scientific and technical. Firstly, in the regions of interest, such as in the LEO environment, electrons of high enough energies to be detected by these types of detector have vastly higher populations than detectable protons in the same regions. In addition, protons of detectable energies

were relatively easy to measure and the resulting data to deconvolve. The electron images produced are significantly less neat, due to more deflections at the mask and entrance window layers and more perturbations in their paths through the detector. Therefore, techniques usable for electron sources would be expected to be successful for proton or alpha particle sources also.

4.3.1 Analysis procedure

For a given source energy and angular distribution and instrument geometry, a large number of particles were flown and the list of locations and energies deposited in each interaction in the active region was saved in a simple CSV datafile. Then the data relating to the number of particles needed for the analysis was read into memory by a Matlab script. All further data processing and presentation was done in Matlab.

Simulated detector readout

Once the list of full interaction data was stored in a Matlab array, it was necessary to convert this into a possible image frame. For each interaction in the active region or scoring volume, the position and energy deposited was read in. The energy was then accumulated in one of 256 by 256 pixel bins, depending on the x and y positions.

In this process, several simplifications were made. Firstly, the detector was treated as though it was 100% percent efficiency in the active region and no interactions were detected in the assumed dead layer. In reality, the charge collection efficiency (CCE) may depend on interaction depth or position within a pixel. For any scaling of the complete CCE, the technique is unchanged: the expected energy deposition by particles at a certain energy would be determined by experiment, and the result would be a scaling of the final deconvolved image. Calibration of the detector and the instrument is needed to account for this effect entirely, but the simulation of the technique is independent of the scaling.

There is also no account taken of potential charge spreading, whereas in reality this would depend on the detector used and the incident particle energy. It is also the case that with the detectors considered, the resolution of the detector is several times smaller than the resolution of the mask, so the loss of signal incurred by small-scale charge spreading would be relatively minor.

In addition, there is no simulation of a digitisation step or other electronics behaviour. All energy deposited in the region of one pixel is summed. Again, how realistic this is depends on the detector and its readout electronics.

The result is related to a pixellated version of the projected image at the detector plane, but takes into account the interactions in and around the dead layer. These interactions can cause significant deflections of particles, especially electrons, and even cause them to fail to reach the active region at all.

Deconvolution

To prepare for the deconvolution step, the mask shape was created in the same orientation as the orientation in the simulation. For the URA designs, these had been initially created in Matlab and exported to the Geant4 simulation. In the case of the random mask, the mask was created in the Geant4 simulation simply by choosing a value of either 0 or 1 at each grid position. This was then exported to a binary file and read into Matlab. In each case, the same array for the masks was available in both Matlab and Geant4.

This array then needed to be resized to be on the same physical scale as the detector. This can be done using various sampling methods, but for these simulations, a nearest neighbour resizing algorithm was used. In addition, the pixel and mask element sizes were chosen so that the mask image when resized to the scale of the detector had all elements the same size as each other.

The mask array was then adjusted so that the values representing holes and mask elements produced a mean value of 0 over the whole mask. This avoids false gradients towards the edge of the reconstructed source data.

The deconvolution step, such as cross-correlation, was then performed, which acts as a transformation into angular space. The pixels in the resulting image represent a solid angle based on the resolution and scale of the detector and mask images. This is then straightforward to map onto a grid of angles.

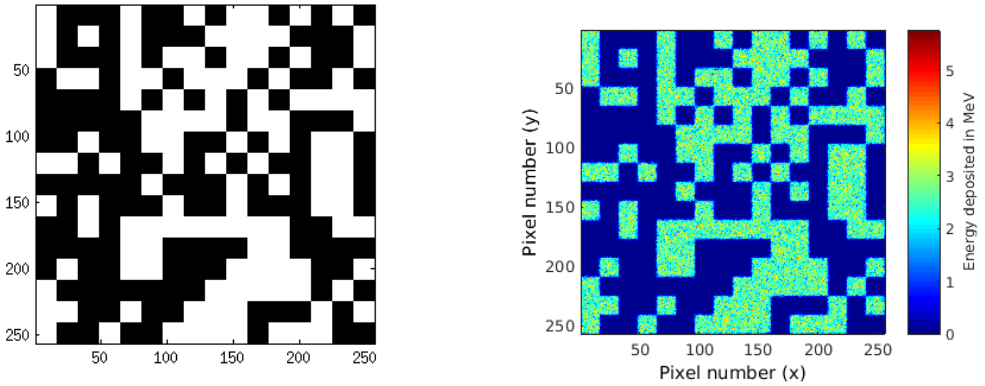
As an example of the process of transformation between the Medipix readout frame and the deconvolved reconstruction in angular space, a simulation is described here. This simulation used a detector with Medipix dimensions and pixel size, in conjunction with a 16×16 random mask with approximately a 50% open fraction with the same dimensions as the detector. This was run with a unidirectional source of 200 keV electrons to make clear the shape of the mask on the detector. A total of one million particles were flown for this simulation, aimed at and around the detector at a square target area of 15 mm×15 mm.

Both the shape of the random mask scaled to the same size as the detector and the resulting image produced on the Medipix are shown in figure 4.7. For this simple case of a parallel electron flux, the shape of the mask is exceptionally clearly seen on the detector.

A simple cross-correlation was performed of the mask and the simulated Medipix output and the result is shown in figure 4.8; the unidirectional nature of the source is very clearly reconstructed. This is normalised to the known source: the central peak relates to a flux of one million parallel particles over an area of 1.5 mm×1.5 mm and the rest of the FOV is empty. These scaling coefficients can then be used for all other central sources of the same electron energy, or be scaled for energy measurements rather than particle number measurements.

As can be seen in this image, there are still some artifacts remaining in the background. A large fraction of this is because of the shape of the random mask shape which is not optimised for cross-correlation.

For a situation such as this, the Lucy-Richardson algorithm is expected to give a very clear result, which is shown in figure 4.9. The implementation of the algorithm available in Matlab was used with 10 iterations. This was a substantially slower



(a) The random mask

(b) Simulated Medipix output frame

Figure 4.7: The mask shape used for this simulation at the same scale as the Medipix frame, and the energy deposited in each pixel of the active region of the detector. One million 200 keV electrons were flown.

process than cross-correlation, but performed particularly well for this case. The Lucy-Richardson algorithm is especially well-suited to a parallel source. In other cases, good results for a series of very narrow sources may be obtained by using the algorithm to find the brightest points and removing their signal from the data before using the deconvolution for a second time.

This is possibly the simplest case that could be simulated with an electron source. More complicated instrument designs and source distribution simulations are described next.

4.3.2 Energy limits of the instrument

These results have limits of the energies and types of particles that can be measured, largely based on the type of detector, the dead layer thickness, and the active region thickness. The limits of what can be deconvolved depend on the particles that are measured being clearly localised on the detector to a position well within the scale

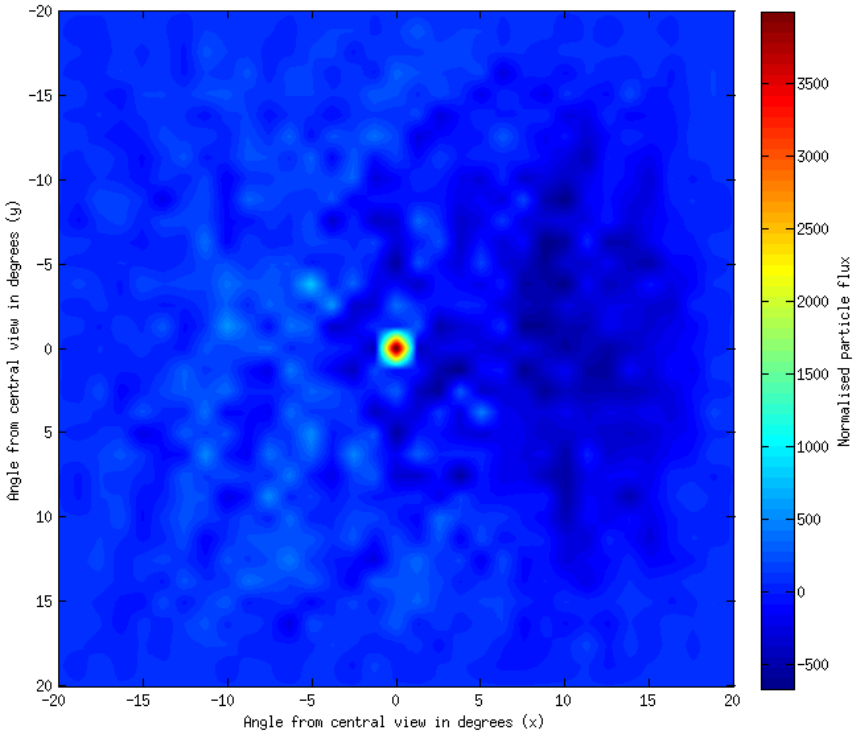


Figure 4.8: The cross-correlation of the arrays in figure 4.7 (a) and (b). This is scaled to the number of particles in each angular bin of 4.7 arcminutes.

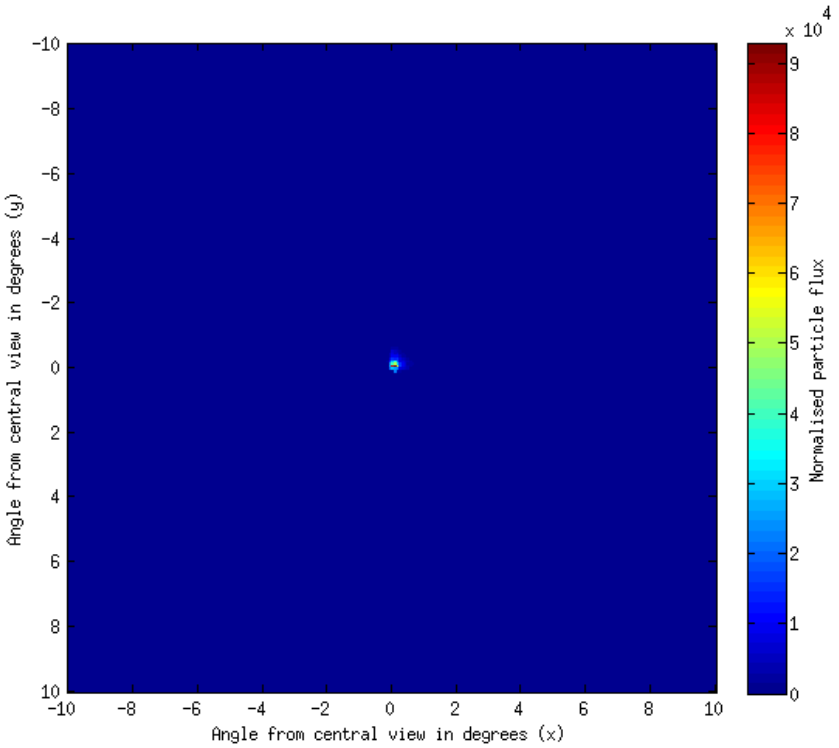


Figure 4.9: The deconvolution of the arrays in figure 4.7 (a) and (b) using a Lucy-Richardson algorithm with 10 iterations.

of the mask.

With the possible Medipix-like geometry, the dead layer is taken to be 10 µm of silicon dioxide and the active region to have a thickness of 300 µm of silicon. Using the data tables from ESTAR and PSTAR, the limits for energies that have a range ending in the active region correspond to the range of electrons between 35 keV and 225 keV and protons between 850 keV and 6.25 MeV. Higher energy particles can be detected, but not all the energy will be measured.

Since electrons are capable of creating multi-pixel tracks in the detector, it can be difficult to identify the entry position of an electron for certain energies. As can be seen in figure 4.10, the expected spatial distribution of the deposited energy varies with electron energy. The maximum expected detected track lengths of electrons in this detector geometry are up to around 3 or 4 pixels for electrons with energies between approximately 200 keV to 300 keV. Higher energy electrons are more likely to reach the back of the active region before losing enough energy to be significantly deflected.

The consequences of this distributed signal are a contribution to the spread of the reconstructed signal. However, this distribution is always on a smaller scale than the scale of the holes in the mask so the deconvolution is still possible for all energies within the range selected.

The same simulations have been run for protons in the range of energies that would be detectable with this geometry and are shown in figure 4.11. It is clear that the PSFs produced by protons are almost always contained within the area of a single pixel, because they travel in almost straight lines through silicon. With this geometry, it can be seen that the maximum signal is produced for protons around 6 MeV - as predicted above using the range tables, higher energies pass beyond the active region of the detector.

Using the ESTAR and PSTAR data for the 0.5 mm of tungsten that was used in these simulations, it is expected that electrons below 2 MeV and protons below 20 MeV will be completely blocked by the opaque parts of the mask. Environments

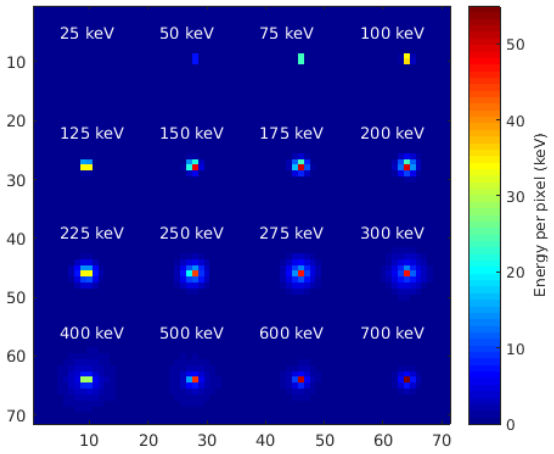


Figure 4.10: The PSF of electrons of different energies in the Medipix-like geometry used for these simulations. 10,000 particles of each energy were flown, and the resulting energy deposition was binned into the detector pixels and re-normalised to the scale of a single particle.

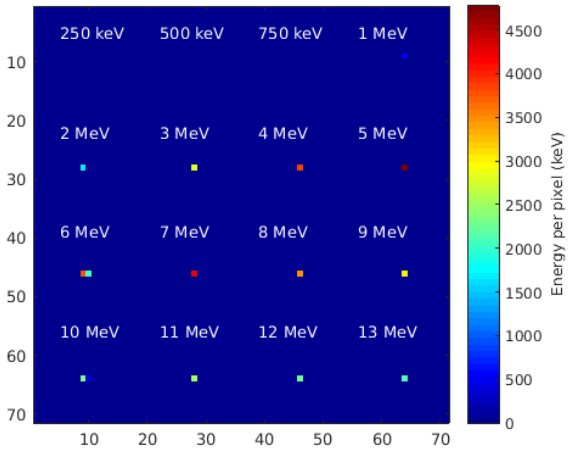


Figure 4.11: The PSF of protons of different energies in the Medipix-like geometry used for these simulations. 10,000 particles of each energy were flown, and the resulting energy deposition was binned into the detector pixels and re-normalised to the scale of a single particle.

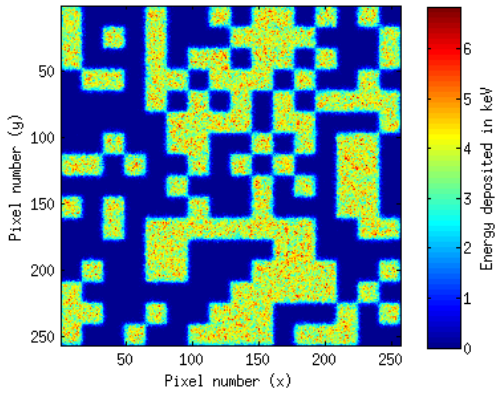


Figure 4.12: Image of energy deposited in each pixel of the active region of the detector for a simulation of a random mask with a source of one million parallel 300 keV electrons. The deconvolution is shown in figure 4.14.

with significant proportions of particles above these energies would receive a higher degree of background signal on the detector and difficulty would be encountered in reconstructing the signal. There is also a risk of particles damaging the detector or any electronics positioned behind it.

Within these energy limits, however, the coded aperture method has been simulated. For example, a slightly higher electron energy in the same simulation as the previous example, using 300 keV electrons rather than 200 keV is shown in figure 4.12 which is visually similar to the Medipix output in figure 4.7 (b). The exact scaling between the two depends on the depth of energy deposition and the pattern of charge collection within the detector, but this simulation appears to produce approximately 50% higher signal level, as would be expected.

The exact scaling between the two depends on the depth of energy deposition and the pattern of charge collection within the detector, but for these energies the 300 keV simulation produces a signal level very close to 50% higher than the 200 keV simulation, as would be expected. The distribution of pixel values in these images is

plotted in figure 4.13, showing that the ratio of average energies deposited per pixel is close to the ratio of the incident electrons.

Visually, the reconstructed distribution shown in figure 4.14 looks very similar to figure 4.8 above.

4.3.3 Number of particles required

The incident flux required to regain a clear signal with a high SNR in the deconvolution is a key measure of the integration time needed in a particular environment to reconstruct the flux. For a specific situation, in this case the unidirectional incident flux of 300 keV described above, the same simulation was run with different numbers of sources to view the improvement to SNR with an increase in signal. The single direction flux was used to demonstrate the behaviour of the output with increasing flux and to represent an upper limit of the reconstruction possible with a certain number of particles.

Four orders of magnitude were used: 100, 1000, 10000 and 100000 particles, flown perpendicular to the mask in a 15 by 15 mm square. The fraction of these that reach the front of the detector through the open half of the mask would therefore be expected to be

$$\frac{14.08^2}{15^2} \times \frac{136}{256} = 0.468$$

where the first part is the area of the detector over the full area of the simulated flux, and the second part is the open fraction of this random mask. This would be scaled by the relevant fraction if smaller holes were used in place of fully-open grid elements.

The simulations in the previous section were run with a total of one million particles. By selecting a smaller number of particles, the limits of the number of particles needed for reconstruction of a point source can be found. The output of the Medipix frames are shown in figure 4.15.

The deconvolution of these Medipix frames in figure 4.16 shows that only a few

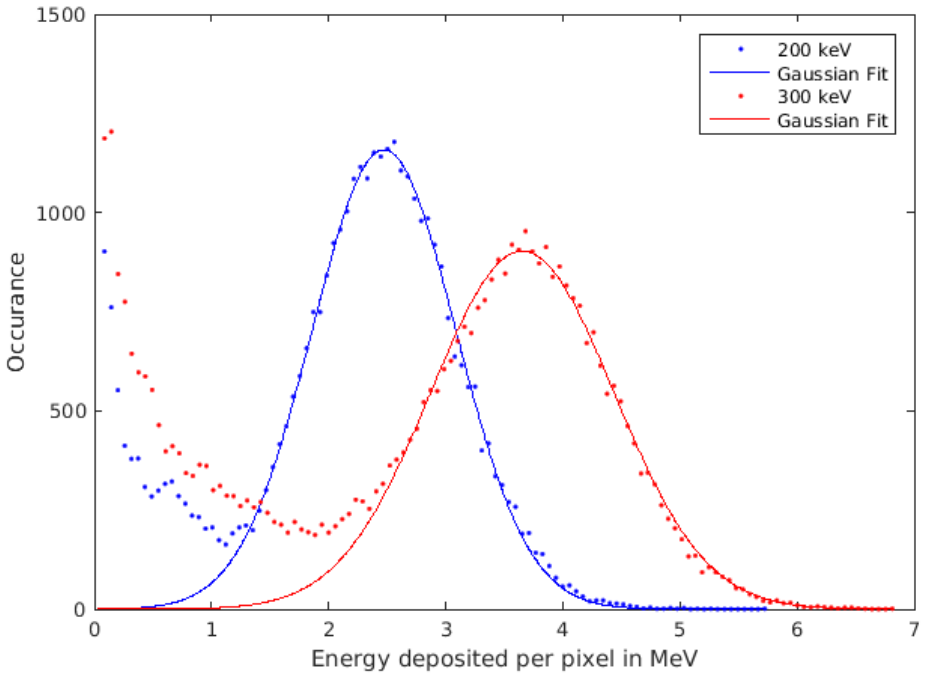


Figure 4.13: The pixel energy distribution of 200 keV and 300 keV electrons using a random mask and a parallel source, with the occurrence calculated per bin width of 58 keV. The raw output frames are shown in figure 4.7 and figure 4.12. A Gaussian curve has been fitted to the main peak. The central energy of this fitted curve is 2.47 MeV for the 200 keV simulation and 3.65 MeV for the 300 keV simulation - a ratio of 1.48.

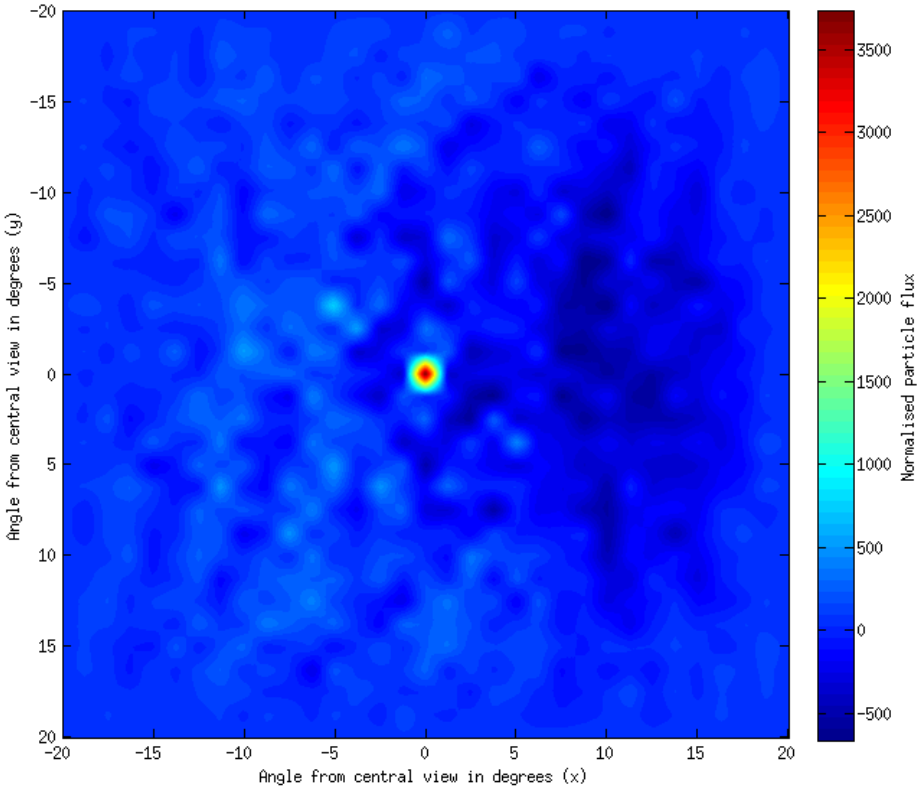
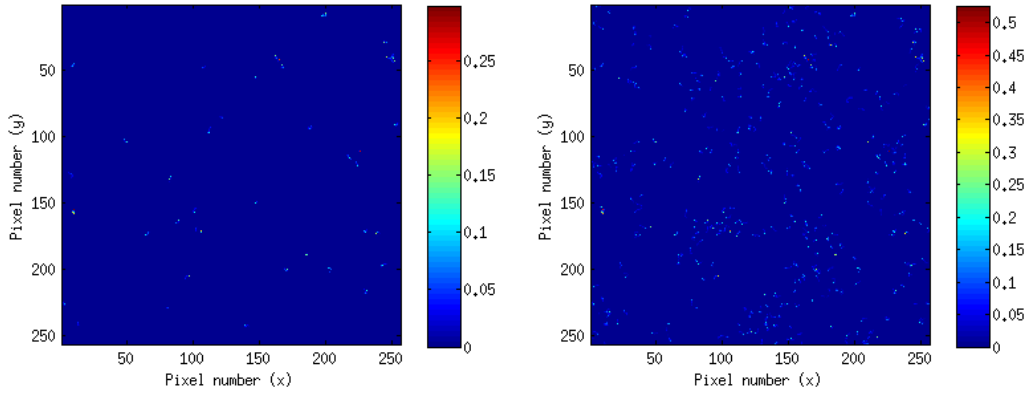
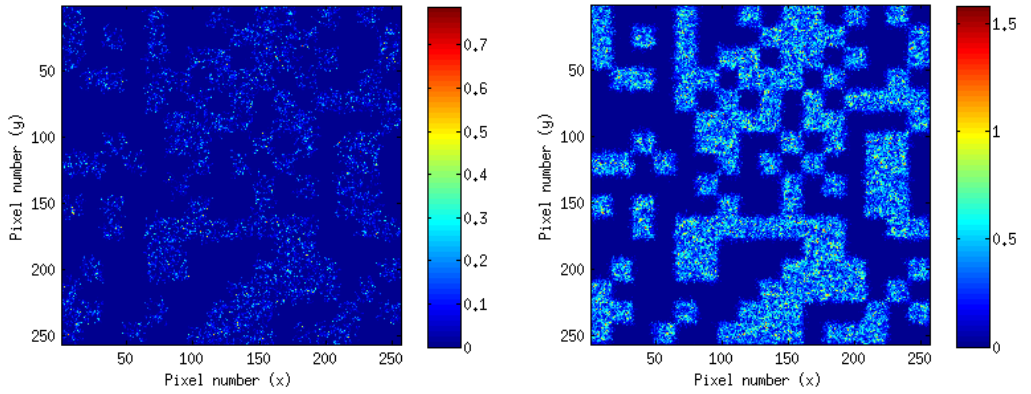


Figure 4.14: The Medipix data shown in figure 4.12 cross-correlated with mask shape and normalised to a number flux of one million particles.



(a) 100 particles

(b) 1000 particles



(c) 10000 particles

(d) 100000 particles

Figure 4.15: Image of energy deposited in each pixel of the active region of the detector by the flux of a range of numbers of 300 keV electrons

Table 4.1: The SNRs for simulations including each number of particles.

Number of particles	100	1000	10000	100000
SNR	11.24	32.63	47.26	43.93

particles are required to produce a very clear reconstructed signal in the case of no background flux or detector noise.

The number of particles required in a particular case is depends heavily on the energy and distribution of the source. However, as long as the pattern is clearly seen on the detector, the deconvolution very quickly converges, since speckle in the image at a small scale is usually averaged out.

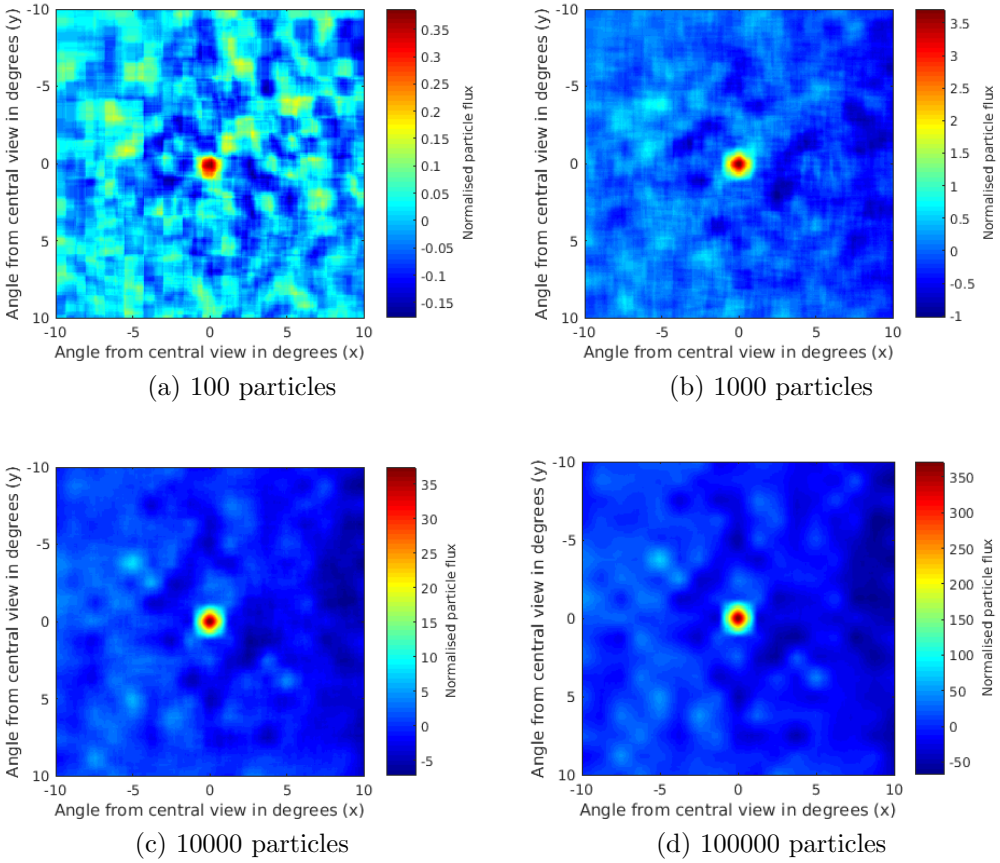
A cross section across the deconvolved outputs shown in figure 4.16 is shown in figure 4.17. They are scaled to the peak height of each cross section for ease of visual comparison.

An estimate of the SNR can be calculated as the ratio of the peak value of the signal to an estimate of the distribution of values in the rest of the image. This noise level is estimated by calculating the standard deviation of an arbitrarily-chosen section of the deconvolved output. In this case, the section chosen was a square of 100×100 pixels where both x and y pixel values were from 101 to 200.

Using this method, the SNRs of the cases shown in figures are given in table 4.1. This dependence of SNR on number of particles in this case is given further in figure 4.18. A million particles were flown, and the particle number shown represents the inclusion of that number of particles from the beginning of the list of interactions. In this case, the SNR can be seen to increase up to a number flux of around a few thousand particles.

4.3.4 Mask shape choice

The previous examples were all run with the same 16×16 random mask. It was expected that this could be significantly improved upon with the use of a mathematically-



(a) 100 particles

(b) 1000 particles

(c) 10000 particles

(d) 100000 particles

Figure 4.16: Simulation data from the Medipix data shown in figure 4.15 deconvolved by cross-correlation.

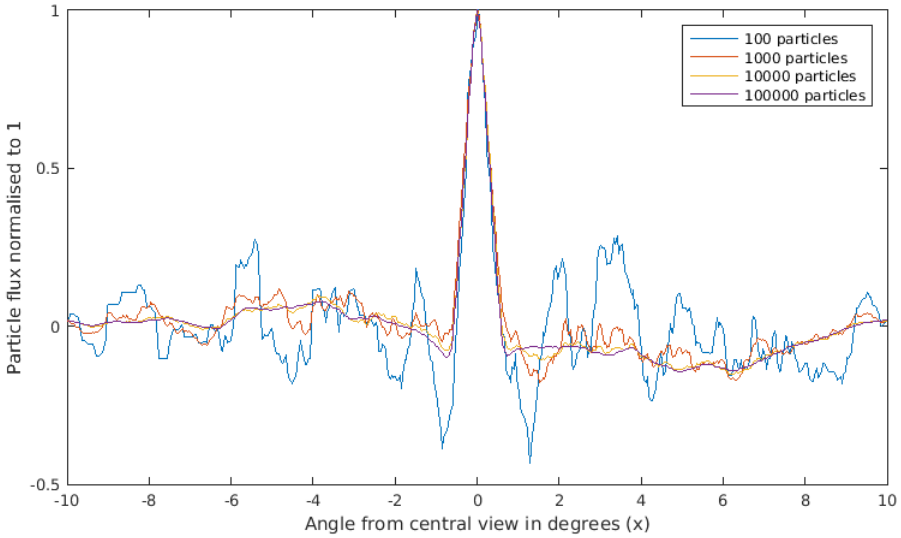


Figure 4.17: Cross section from simulation data from the Medipix data shown in figure 4.15 deconvolved by cross-correlation. The cross sections are scaled so the peak height in each case is 1.

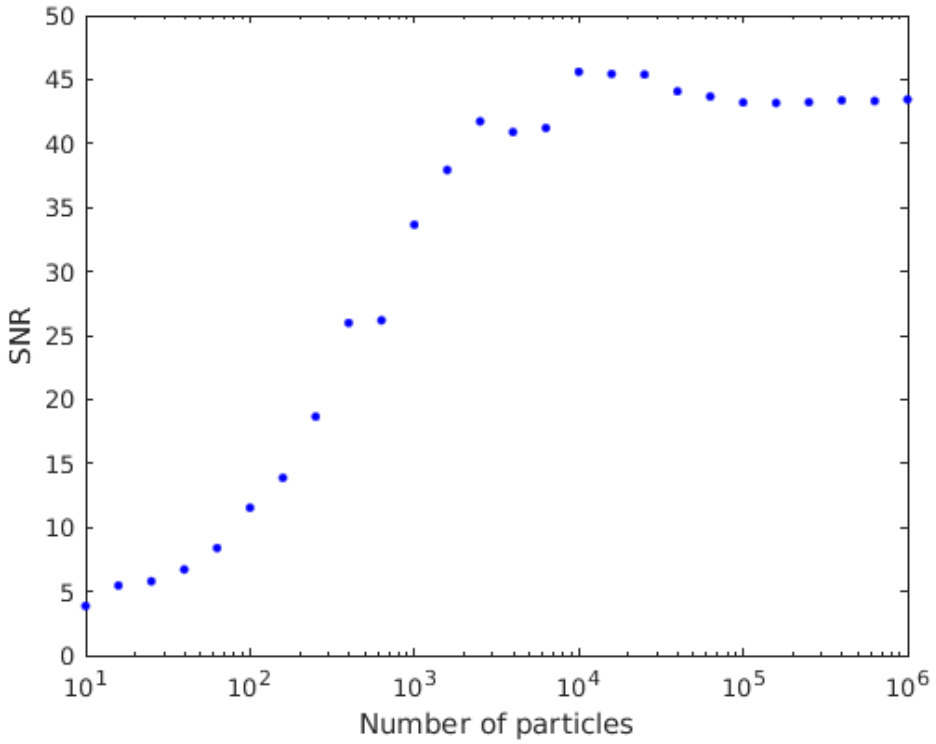
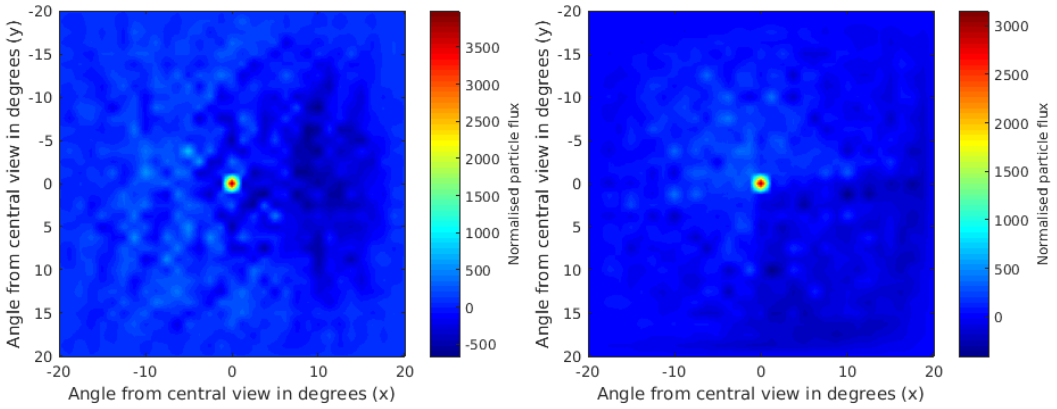


Figure 4.18: The SNRs for simulations against the number of particles counted in the simulation for the geometry given in the results shown in figure 4.16



(a) Random mask

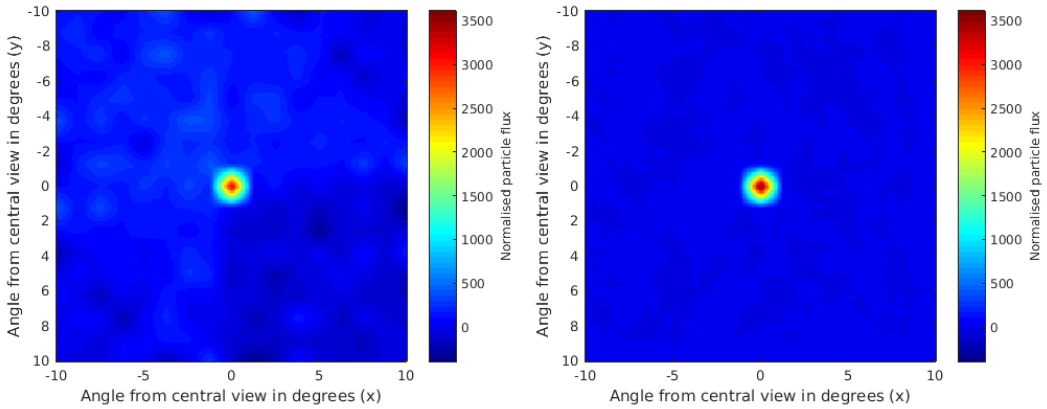
(b) PBA

Figure 4.19: The cross-correlation deconvolution of a random mask, and a 16×16 PBA, with a 200 keV electron beam source

superior URA.

Several URAs were simulated. The results from simulations using both a random mask and the 16×16 PBA scaled to the detector size are given in figure 4.19. The background of the deconvolved image, which would ideally contain no signal, has significantly fewer artifacts in the image which used the PBA. The PBA simulation produces an SNR of 60.3 while the random mask produces an SNR of 46.4; both were calculated using the method given in section 4.3.3.

The URAs were, as expected, better performing, so one of these was decided upon. The convenience afforded by selecting a mask shape which matched the shape of the detector led to the selection, and therefore PBA-based mask shape was chosen. The rest of the simulations were run using a 16×16 PBA pattern.



(a) Single mask (b) Doubled mask

Figure 4.20: The central portion of the cross-correlated data of (a) one PBA mask and (b) a repeated PBA mask using a unidirectional source of 200 keV electrons.

4.3.5 Repeating the mask

The previous examples were all run with a single mask shape of the same size as the detector, which significantly restricts the FOV of the instrument. Repeating the mask shape in each direction was suggested by Fenimore and Cannon [1978] as a way to increase the FOV and flux and retain the accuracy of the mask and method. In the case of the design used here, this repetition increases the FOV from approximately 0.5 sr to 2 sr.

The behaviour of a simple simulation with a unidirectional electron flux is shown in figure 4.20. The FOV is larger with a repeated mask, and the artifacts in the central angular region are significantly reduced, since now the FCFOV is over 20° wider. However, any off-axis sources would produce artifacts in this region.

With a single mask, in this case the SNR is 48.8, and with the repeated mask the SNR is 529.8. This almost noise-free background is due to the auto-correlation properties of the PBA: this central region is doubly-coded rather than singly-coded,

so the artifacts created by the deconvolution are reduced.

When a single instance of the mask is used, off-axis sources have a reduced effect on the output frame from the detector. Only a source in the exact centre of the field of view is fully coded. Repeated the mask allows a fully-coded angular region surrounded by a partially coded region.

The rest of the simulations were run with a repeated mask with a total aperture area of 28.16 mm×28.16 mm, or 4 times the detector area.

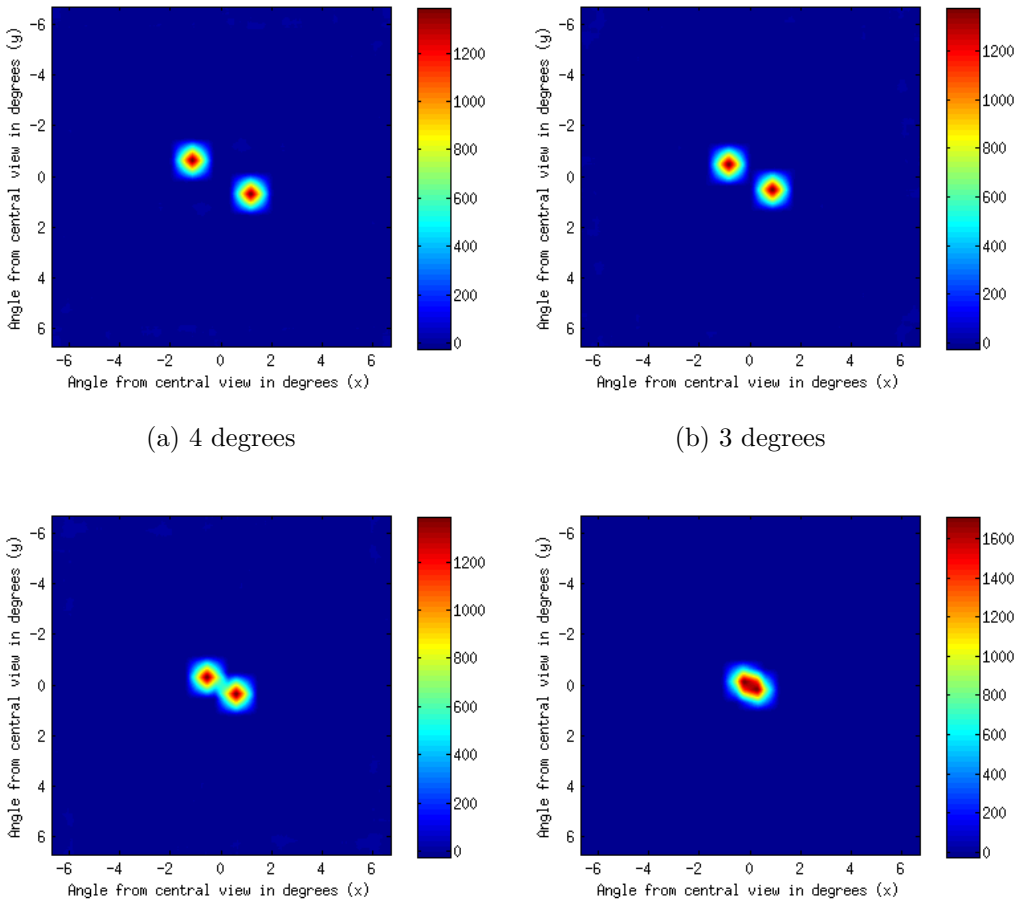
4.3.6 Differentiation of two parallel fluxes

In other studies, for example Sambo et al. [2009], a comparison of the results from simulated data from two distant point sources has been used: the minimum angular separation at which the two sources can be resolved is used as a measure of the capability of the design.

With no flux background from other angles or detector noise, the cross-correlated deconvolution from one million particles is shown in figure 4.21, for several separation angles.

It can be seen that in this case 2° separation can be easily distinguished, while it is difficult to resolve 1° separation into two separate sources. This estimate of resolution is useful as a comparison between instrument designs, but the angular resolution possible is highly dependant on the exact source distribution, and also on the radiation background and the noise in the detector.

These images are shown as a slice across the image directly through the peaks of the sources in figure 4.22. The 2° separation case can be clearly distinguished with two separate peaks with a dip down to 50% below the peak level between the two. Closer spaced peaks would be difficult to distinguish.



(a) 4 degrees

(b) 3 degrees

(c) 2 degrees

(d) 1 degree

Figure 4.21: The deconvolved results of 200 keV electron parallel beam simulations with no noise. In each case, one million particles were flown. The flux is normalised to the same degree as in figure 4.8.

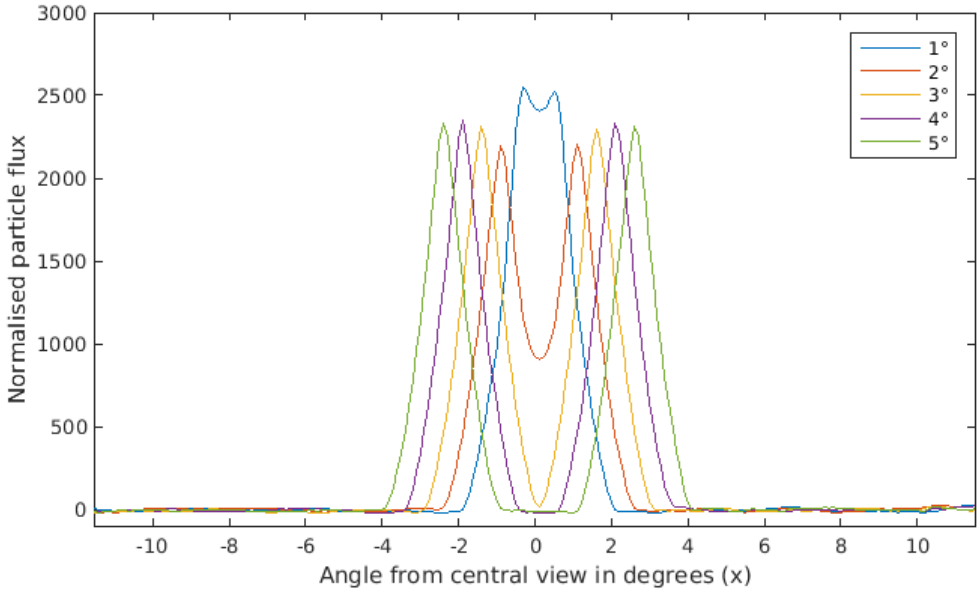


Figure 4.22: A slice through the two dimensional intensity plots in figure 4.21 with the addition of the case with a further separation distance of 5°.

4.3.7 Sources with a particle background distribution

Two different effects were examined here: a background radiation field, and the randomly-distributed noise that might be expected as thermal noise or dark current from within a detector. Each of these would add more complication to the Medipix frame and make an accurate angular reconstruction more complicated.

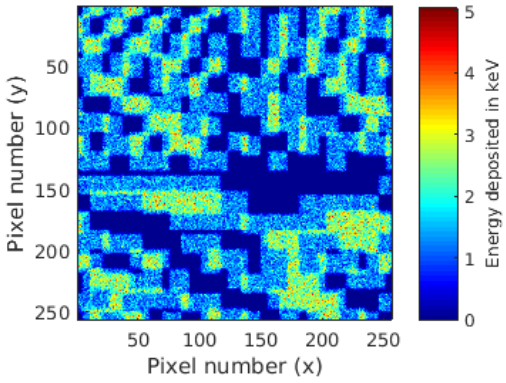
The data from the 2° separation simulation shown in figure 4.21 (c) was used but at the Medipix stage further data were added. A simulation was run of both one million and ten million 200 keV electrons evenly distributed over a range of angles from −20° to 20° in both the x and y directions. This would represent an isotropic particle flux distribution. The sum of these two simulations in the detector is shown in figure 4.23 (b).

In addition, the effect of detector noise was observed. This might be the effect of dark current in a detector or some other random effect. A field of Gaussian noise with a standard deviation equivalent to 5 keV deposited in the detector was used. The sum of this field with figure 4.23 (b) is given in figure 4.23 (c).

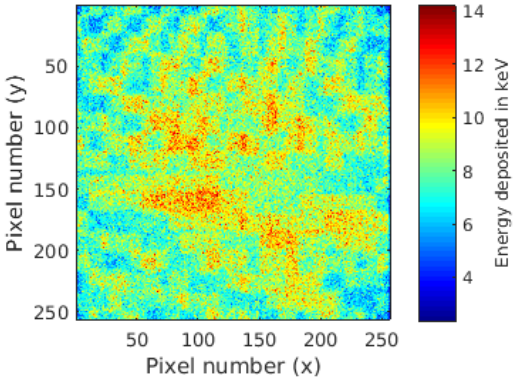
The deconvolution of these detector frames is shown in figure 4.24. In this image, (a) is both background- and noise-free. The first column is noise free, and the second column contains the noise field described above. The first row contains no isotropic particle flux, the middle row has one million background particles (the same number as the number in the signal) and the bottom row has ten million background particles.

There is an increase in the signal at the edges of the cross-correlated images from the simulations with a significant proportion of isotropic flux. A cross-section of these along a line passing through both peaks is shown in figure 4.25. This increase at the edges is largely due to the collimation effect of the holes in the mask. These filter out some of the particles which are incident off-axis.

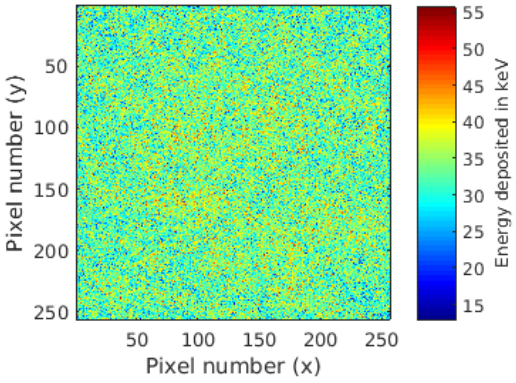
Noise in the detector is shown to result in some noise in the deconvolved image, while an isotropic flux increases the background in the recovered image along with additional signal levels at the edges. Picking strong sources out from such back-



(a) Signal

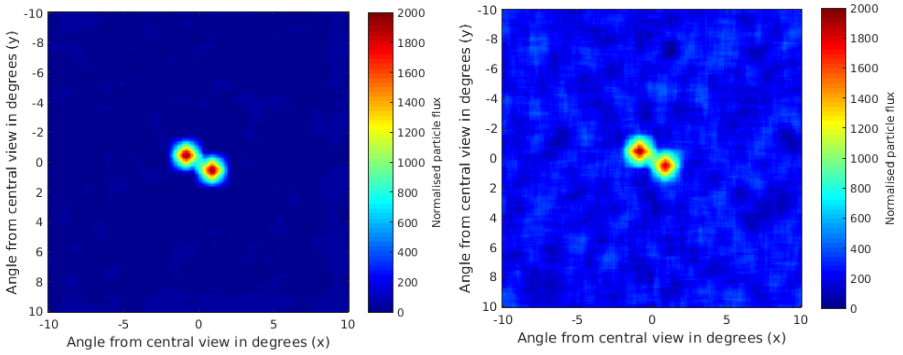


(b) With 10 million background



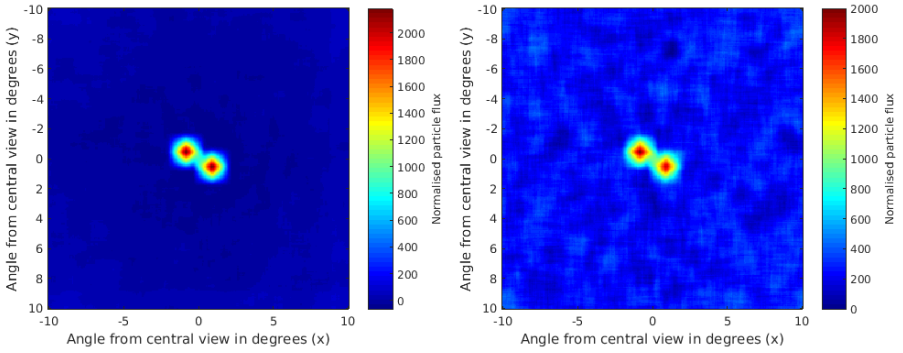
(c) With background and noise

Figure 4.23: The signal on the detector with additional effects. (a) shows the signal with no background or noise(b) shows the additional of an evenly-distributed background flux. (c) shows the signal and background flux with the addition of a strong noise field in the detector.



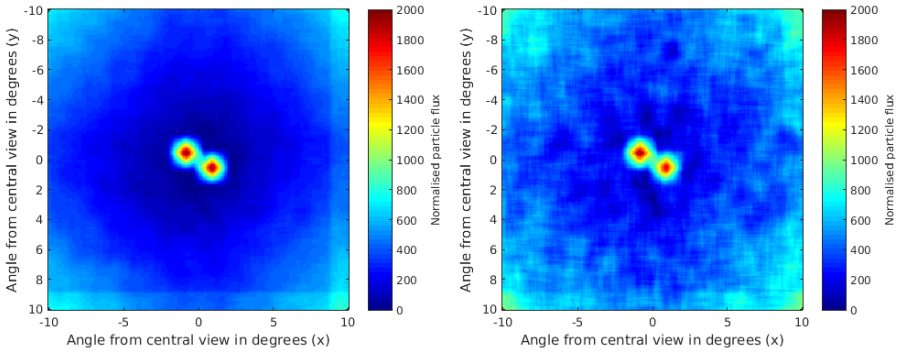
(a) no background, no noise

(b) no background, noise



(c) 1 million background, no noise

(d) 1 million background, noise



(e) 10 million background, no noise

(f) 10 million background, noise

Figure 4.24: The deconvolved results of one million 200 keV parallel electron simulations with 2° separation with noise and background. In each case, one million particles were flown.

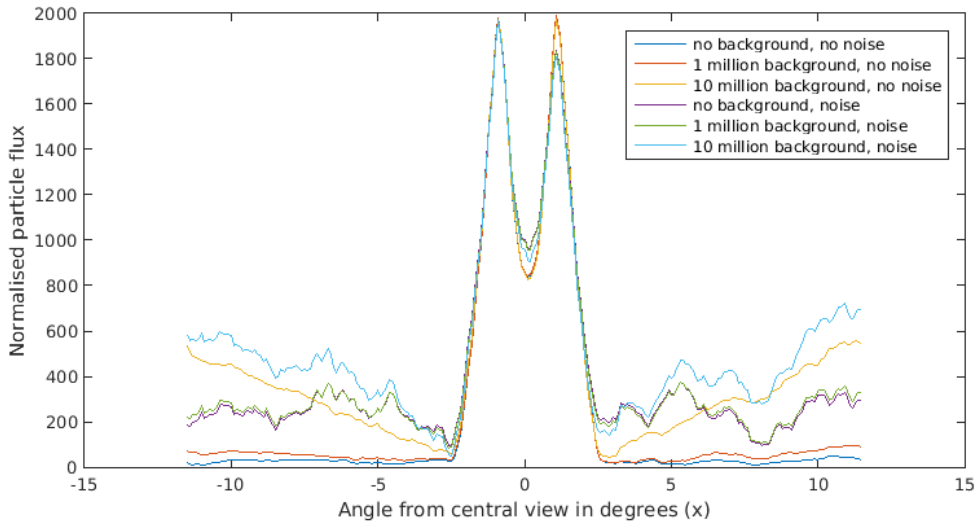


Figure 4.25: The cross section through both peaks for each of the recovered images shown in figure 4.24.

ground and noise sources is possible, even when the noise levels are high enough that structures on the frame cannot be seen by eye. This demonstrates that the deconvolution is very insensitive to detector noise. This noise should, however be estimated and taken into account for a certain exposure time and environment. In addition, a clear signal can be seen in a background distribution. When there are high signals at the edges of the field of view, the deconvolution suffers from increased signal at the edge of the image frame

4.3.8 Measuring extended sources

In reality, it would be expected that realistic distributions in space would have a finite width. Regions such as the aurora and electron strahl in the solar wind can be strongly aligned, usually to angles within a plane.

This simulation was of incident flux of 20° length in one direction and varying

width in the perpendicular direction. This 'pancake' distribution was aligned at an angle of 30° from the y axis of the instrument, to represent an arbitrary angle. The results of this reconstruction for angular widths of 1° to 4° using 200 keV electrons are shown in figure 4.26.

It can be seen that at least in these cases, the flux width and strength can be reconstructed well. The cross-sections of these images will be shown in figure 4.28.

These simulations are robust against detector noise. A distributed isotropic background with a total of one million particles, or a flux of $5.05 \times 10^7\,\mathrm{sr}^{-1}\mathrm{cm}^{-2}$, was added to the simulations. The deconvolutions of these simulations are shown in figure 4.27.

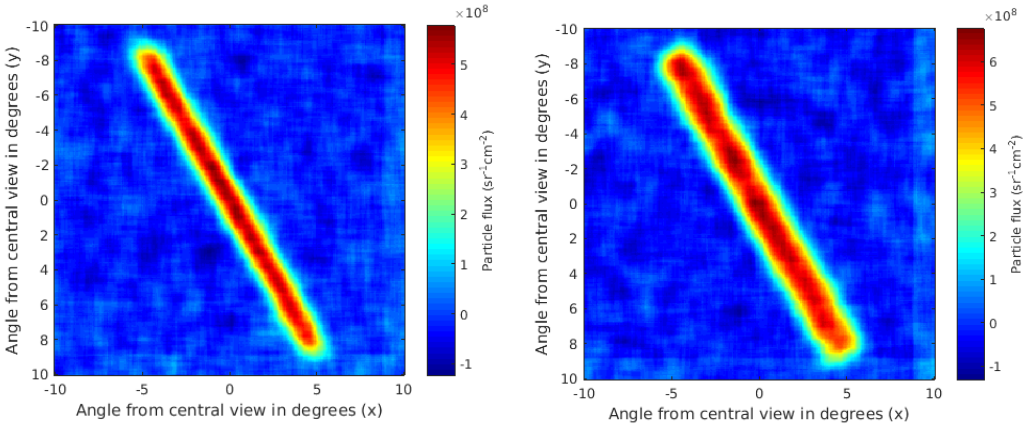
A cross section across the centre of the extended source for each of the recovered images shown in figure 4.26 and figure 4.27 is shown in figure 4.28. The width of the source can be seen, and the background distribution adds an offset, in particular at the edges of the image. The FWHMs of the 1°, 2°, 3° and 4° cases without background are 1.48°, 2.07°, 3.01° and 3.94°.

The increase at the edges due to collimation can be somewhat mitigated by modifying the shape of the mask image when deconvolving these parts of the image, but the degree to which this helps changes with the exact source distribution. Further optimisation of the deconvolution algorithm may improve these retrievals.

4.4 Simulation conclusions

These initial simulations demonstrate that a coded mask instrument is capable of high-resolution angular measurements of point and extended angular sources. From these simulations, a good choice can be made for mask shape, size and thickness and geometry of the instrument, as well as the data collection, reconstruction and analysis plans. In addition, the exact geometry could be modified for a different radiation environment or scientific application.

Based on these simulations, this instrument design would be capable of differen-



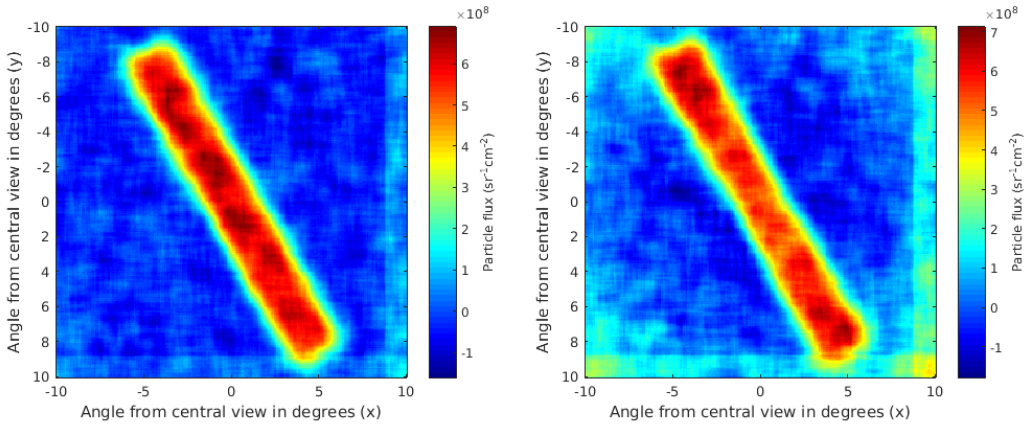
(a) 1 degree

(b) 2 degrees

(c) 3 degrees

(d) 4 degrees

Figure 4.26: The deconvolved results of 200 keV electron distributions, using the same normalisation as in the earlier simulations.



(a) no background particles

(b) 1 million background particles

Figure 4.27: The deconvolved results of 200 keV electron distribution of 3° width with background flux compared to the background-free frame shown in figure 4.26.

tiating angles at an order of 1° resolution for electrons above approximately 40 keV and protons above around 900 keV, and is highly resilient to noise in the detector.

Once the proof-of-concept version of the instrument was tested, as described in chapter 5, an instrument based on this simulated design was simulated in space-like radiation environments as detailed in chapter 6.

4.4.1 List of simulated scenarios

Table 4.2 shows the most important simulations run to compile the results in this chapter. Further simulations were completed, but not quantitatively analysed.

4.4.2 Table of results

A summary of the results is shown in table 4.3 giving the SNR of the simulations that have been shown in this section.

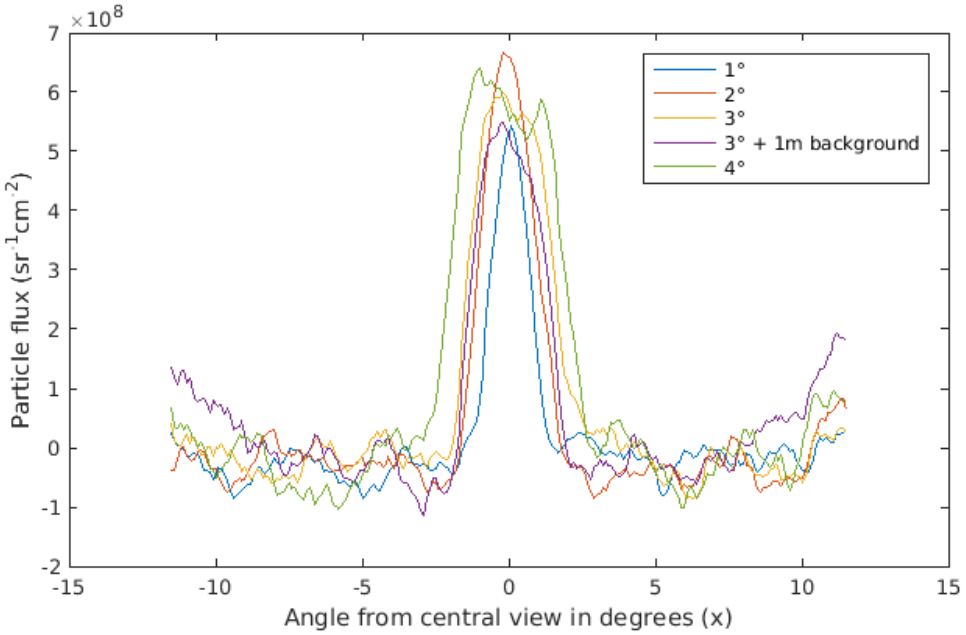


Figure 4.28: The cross section across the centre of the extended source for each of the recovered images shown in figure 4.26 and figure 4.27.

Table 4.2: A summary list of the main simulation situations used for this chapter.

Mask shape	**Repetition**	**Scale**	**Particle**	**Energy**	**Distribution**
Random 16	single	880 um	Electron	100 keV	Monodirectional, perpendicular
				200 keV	Monodirectional, perpendicular
				300 keV	Monodirectional, perpendicular
				200 keV	Bidirectional, 5° separation
			Proton	1 MeV	Monodirectional, perpendicular
				2 MeV	Monodirectional, perpendicular
				3 MeV	Monodirectional, perpendicular
	Repeated 2x2	880 um	Electron	100 keV	Monodirectional, perpendicular
			Proton	1 MeV	Monodirectional, perpendicular
PBA 16	Single	880um	Electron	100 keV	Monodirectional, perpendicular
				200 keV	Monodirectional, perpendicular
				300 keV	Monodirectional, perpendicular
			Proton	1 MeV	Monodirectional, perpendicular
				2 MeV	Monodirectional, perpendicular
				3 MeV	Monodirectional, perpendicular
	Repeated 2x2	880 um	Electron	100 keV	Monodirectional, perpendicular
				200 keV	Monodirectional, perpendicular
				100 keV	Bidirectional, 5° separation
				200 keV	Bidirectional, 5° separation
				300 keV	Bidirectional, 5° separation
				200 keV	Bidirectional, 1° separation
				200 keV	Bidirectional, 2° separation
				200 keV	Bidirectional, 3° separation
				200 keV	Bidirectional, 4° separation
				200 keV	Omnidirectional: x/y ±20°
				200 keV	Broad line: x ±10°, y ±0.5°
				200 keV	Broad line: x ±10°, y ±1.0°
				200 keV	Broad line: x ±10°, y ±1.5°
				200 keV	Broad line: x ±10°, y ±2.0°
				200 keV	Broad line: x ±20°, y ±2.5°
				200 keV	Broad line: x ±12°, y ±1.0°
				200 keV	Broad line: x ±20°, y ±0.5°
				200 keV	Broad line: x ±20°, y ±1.0°
				200 keV	Broad line: x ±20°, y ±2.5°
			Proton	1 MeV	Monodirectional, perpendicular
				2 MeV	Monodirectional, perpendicular
				1 MeV	Bidirectional, 5° separation
				2 MeV	Bidirectional, 5° separation
				3 MeV	Bidirectional, 5° separation
URA 41,43	Single	343 um	Electron	100 keV	Monodirectional, perpendicular
			Proton	1 MeV	Monodirectional, perpendicular

Table 4.3: Table showing the SNRs of the simulated results from the previous sections.

Simulation description	**Figure number**	**SNR**
Parallel beam, 100 keV electrons		
100 particles, random mask	Figure 4.16	11.2
1000 particles, random mask	Figure 4.16	32.6
10000 particles, random mask	Figure 4.16	47.3
100000 particles, random mask	Figure 4.16	43.9
1000000 particles, random mask	Figure 4.14	43.5
Parallel beam, 200 keV electrons		
10 million particles, 16×16 random mask	Figure 4.8	46.4
10 million particles, 16×16 PBA mask	Figure 4.19	60.2
Parallel beam, 300 keV electrons		
1 million particles, 16×16 PBA mask	Figure 4.20	59.5
1 million particles, 2×2 16×16 PBA mask	Figure 4.20	464.4
Bidirectional, 10 million 200 keV electrons:		
1° separation	Figure 4.21	336.9
2° separation	Figure 4.21	233.9
3° separation	Figure 4.21	330.5
4° separation	Figure 4.21	274.9
2° separation, no background, no noise	Figure 4.24	233.9
2° separation, no background, Gaussian noise	Figure 4.24	31.8
2° separation, $5.05 \times 10^7\,\mathrm{sr^{-1}cm^{-2}}$ isotropic background flux, no noise	Figure 4.24	192.9
2° separation, $5.05 \times 10^7\,\mathrm{sr^{-1}cm^{-2}}$ isotropic background flux, Gaussian noise	Figure 4.24	31.2
2° separation, $5.05 \times 10^8\,\mathrm{sr^{-1}cm^{-2}}$ isotropic background flux, no noise	Figure 4.24	27.6
2° separation, $5.05 \times 10^8\,\mathrm{sr^{-1}cm^{-2}}$ isotropic background flux, Gaussian noise	Figure 4.24	18.6
20° band source, 200 keV electrons, flux $5.2 \times 10^8\,\mathrm{sr^{-1}cm^{-2}}$		
1° width	Figure 4.26	29
2° width	Figure 4.26	25.4
3° width	Figure 4.26	22.2
4° width	Figure 4.26	23
3° width, $0\,\mathrm{sr^{-1}cm^{-2}}$ isotropic background flux	Figure 4.27	22.2
3° width, $5.05 \times 10^7\,\mathrm{sr^{-1}cm^{-2}}$ isotropic background flux	Figure 4.27	17

The SNR was calculated for each simulation used in this chapter. For single or double parallel beam simulations, the signal was taken to be the height of the peak above the mean of the background. For the extended sources, the signal was taken to be the mean of an 11×11 pixel square at the centre of the deconvolved image. In the case of the simulations with a background distributed source, the signal was taken to be how far the central signal stood above the background. The noise level was estimated by calculating the standard deviation of a 100×100 section of the background of the deconvolved image. The SNR is then the ratio between these two values and is a good measure of the significance of the noise in the image.

It can be seen that the best and clearest reconstructions of angular distribution, based on the SNR, are those where the signal is one or two parallel beams and the flux is high enough to gain sufficient statistics. The SNR is significantly lower when the source is slightly more distributed, as in the case of the 20° degree long distributed sources. The parallel sources produced SNRs up to around an order of magnitude higher than the extended sources.

The introduction of an isotropic background flux caused a significant reduction in the SNR. For example, the introduction of a flux at 10% of the level of the distributed signal reduced the SNR of the simulation of the two parallel sources at 2° separation from 233.9 to 192.9.

The introduction of simulated white Gaussian-distributed noise in the detector in the three simulated cases also made a significant impact on the SNR value, in some cases close to an order of magnitude. However, the signal was still clearly visible by eye, and the lack of structure caused by random detector noise means that to some degree it would be possible to smooth it out of the image, depending on the scales of the structures that the instrument expects to retrieve. In addition, the noise levels in this simulation were unrealistically high for most detectors, with a standard deviation of the equivalent of 5 keV deposited per pixel, and with the understanding of each pixel's expected behaviour, the detector noise would be expected to be significantly lower in a flight instrument.

Table 4.4: Design and components of the proposed simulated instrument

Parameter	**Value**
Detector geometry	Medipix-like
Mask pattern	16×16 PBA repeated
Mask material	Tungsten
Thickness of mask	0.5 mm
Mask size	28.16 mm
Mask open fraction	50%
Distance detector to mask	4 cm

The geometry of the instrument makes a measurable difference to the capabilities of reconstruction, in particular the choice of a mathematically well-designed mask is shown in the parallel flux simulation case, where a move from a random mask to a PBA increased the SNR from 46.4 to 60.2.

In conclusion, the particular strength of the coded aperture technique for particle measurements is when a source is well-defined in structure - especially parallel or highly collimated beams or 'pancake' distributions. The ability to recover more distributed sources is significantly weaker. The effect of sources which are significantly off-axis on the distortion and artifacts in the image is a concern. Further research may be able to develop geometries or improved algorithms less sensitive to the collimation at the edges of the recovered images.

4.4.3 Parameters of the simulated design

The parameters of the design that has been simulated in this section are summarised in table 4.4.

This geometry can be calculated to give the estimated capabilities shown in table 4.5. In addition to the FCFOV and PCFOV, the influence of angle on flux reaching the detector can be calculated in more detail, as shown in figure 4.29.

Table 4.5: Estimates of the capabilities of such a design

Parameter	**Value**
FCFOV	20°
PCFOV	56°
Theoretical upper limit of resolution	5 arcminutes
Point-source differentiation	2°
Energy ranges	120 keV to 1.25 MeV electrons 700 keV to 18 MeV protons

This represents the number of particles from a given angle that reach the aperture area, travel through a hole and hit the detector, relative to the number of particles travelling normal to the mask reaching the detector. This is independent in each of the x and y directions, since the holes are square. For this mask, where the holes are connected, this is only an estimate - a small amount more flux may reach the detector, although only at parts of the mask. This type of design needs more processing power for accurate deconvolution since the shape of the shadow cast is more variable depending on the incident angle of the particles. However, this type of design is impractical to build - a design for a mask that is possible to manufacture will need some support for all the mask elements.

4.4.4 Parameters for the prototype model

Inspired by the geometry used in the simulations, a proof-of-concept model was produced for characterisation in the laboratory. Since many of the components necessary for the simulation setup would not have been possible to acquire or manufacture, available components were chosen which would be able to demonstrate the concept. As a result, the experiments needed to focus on lower energy particle detection because of the particle sources available for testing. Fortunately, a high-end CCD capable of detecting low-energy electrons was available, and the test model

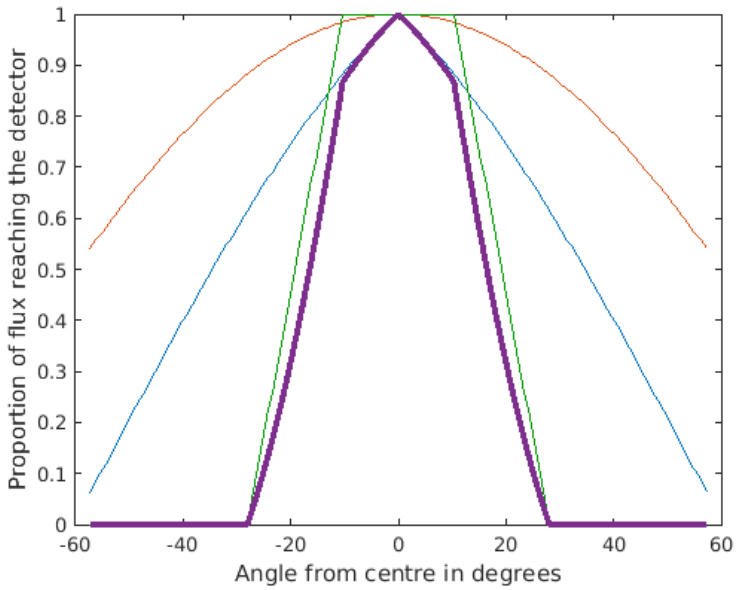


Figure 4.29: The components producing the sensitivity in each of the x and y directions to particle flux. The blue line is scaled to the collimation of the holes in the mask, the orange line is scaled to the effective area of the detector and the green line is the proportion of flux able to pass through the aperture area as it approaches the detector. The thick purple line is the combination of all of these factors and represents the total expected response relative to a normal particle flux.

Table 4.6: Selection of components which are available for use in a prototype instrument

Parameter	**Value**
CCD choice	e2v CCD64
Mask material	Copper-tungsten
Mask pattern	16×16 PBA repeated
Thickness of mask	0.47 mm
Mask open fraction	12.5%
Mask hole size	256 µm
Added CCD layer	None
Distance detector to mask	4 cm

was built around it.

The description of the reasons for the component choices, how an instrument was developed and the results of the experiments will be given in chapter 5. For the purposes of comparison with the simulations here, the main experiment components are summarised in table 4.6.

These values are enough to calculate the expected basic geometrical properties of the instrument. The tungsten-copper pseudo-alloy mask would be expected to shield at least above 750 keV electrons and 15 MeV protons, which is the value for 470 µm of copper. Based on a repeated aperture size, this experimental instrument would have an FCFOV of 12° and a PCFOV of 34°. The detection ratio over flux angle is described in figure 4.30.

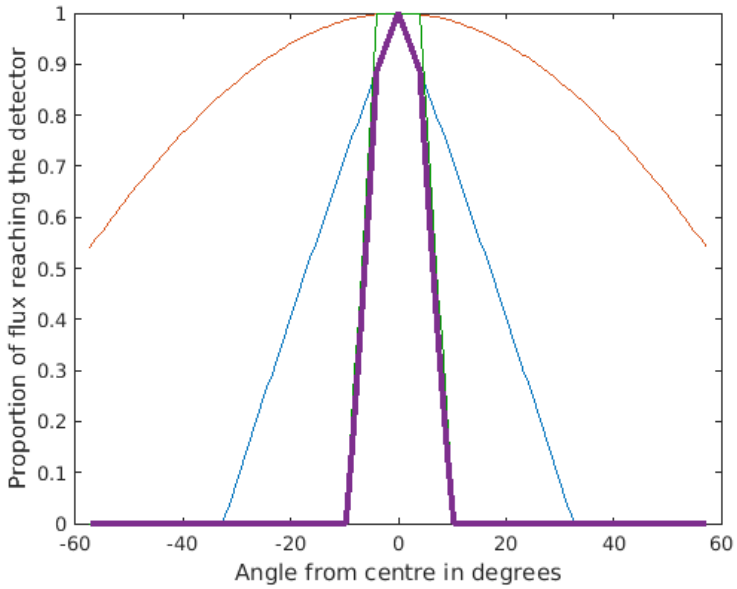


Figure 4.30: The components producing the sensitivity in each of the x and y directions to particle flux. The meaning of each line is as given in the description of figure 4.29 and the thick purple line is the complete particle acceptance at each angle relative to a normal particle flux.

Chapter 5

Experimental setup and results

The experimental setup was intended to demonstrate the potential for the mask to identify the angular distribution of particles, and to verify the simulation setup and environment. Since it was not feasible to reconstruct a space environment entirely, a radioactive beta source was used to provide electrons through a spectrum of energies. With this aim in mind, a simple assembly was conceived, designed and put together, to run a series of tests with various incident electron energy distributions and angular distributions, with the aim of creating a relevant environment to demonstrate the concept of a mask and a silicon detector for electrons. Based on the results that were possible from these initial tests using a samarium source, a second source, carbon-14, was used, and then these were followed by a lower-energy electron source based on an MCP which was capable of producing a higher count rate, at the expense of individual particle identification. These sources allowed a preliminary analysis of the concept in a range of electron environments.

5.1 Concept and requirements

The concept of the experimental setup is shown in figure 5.1. Since resources for the experimental setup production were very limited, the design was heavily con-

strained by the necessity for simplicity and reuse of available parts. The choices, the reasons for them and the implications of these choices are discussed in the following section 5.2.

Figure 5.1 shows the crucial parts to be incorporated into the design: a source of electrons between approximately 20 keV and 500 keV, and a mask and detector setup. A vacuum was necessary for the tests, since any type of atmosphere would absorb and deflect enough particles between the source and the detector that the imaging properties of the instrument would be lost. Since this required the experiments to take place in a vacuum chamber, it was planned to incorporate some means of moving them relative to each other to produce different source angles, avoiding the need to break the vacuum for minor adjustments between experiments.

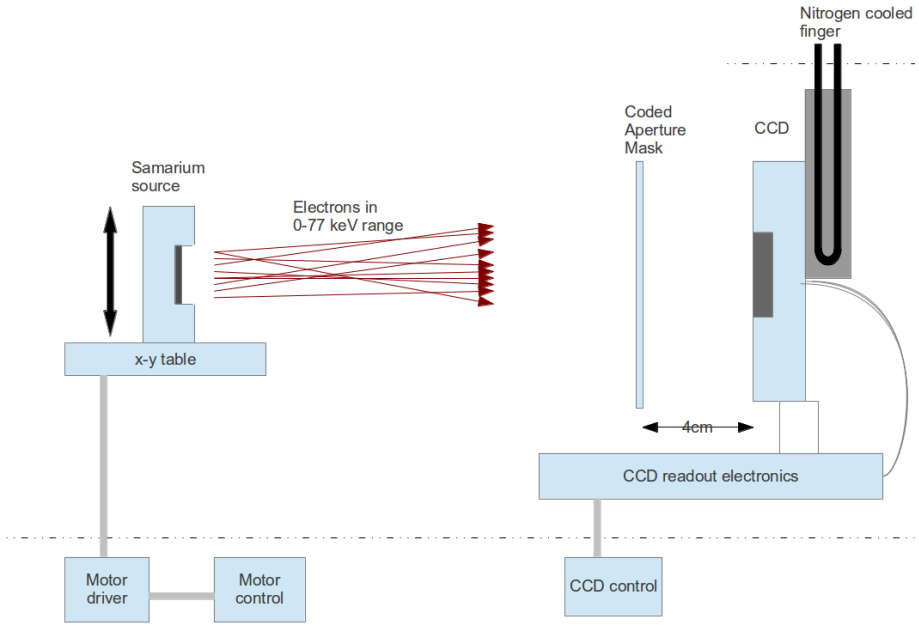


Figure 5.1: A conceptual diagram of the setup to be realised in the laboratory for a samarium source of electrons

Based on this and the limited available components, a setup built around an X-ray CCD from e2v, a copper-tungsten PBA mask and, initially, a radioactive electron source was designed.

The model used for final manufacturing is shown in figure 5.2. This does not include harnessing, which required the minor modification of some of the components.

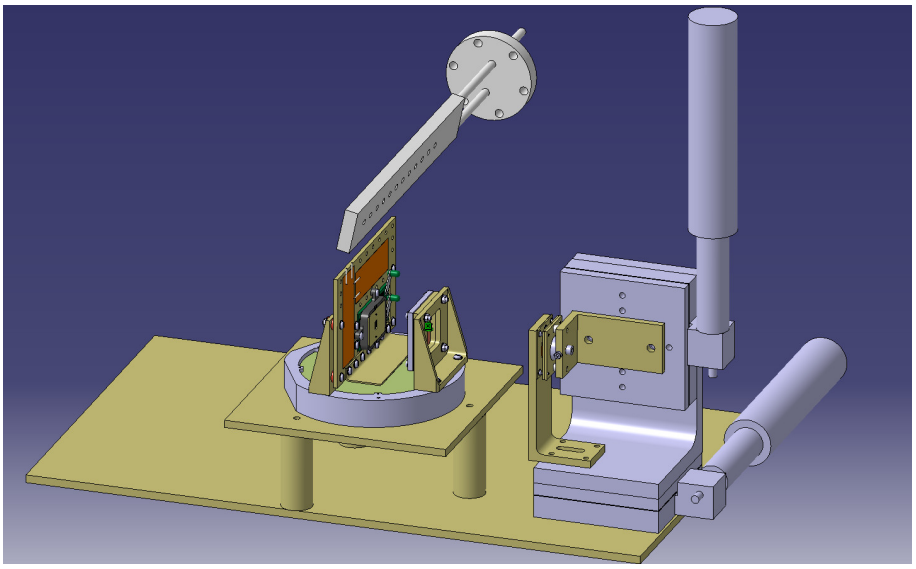


Figure 5.2: CAD of the experimental setup in the vacuum chamber

5.1.1 Facilities

The vacuum requirements for the setup were not particularly demanding; the greatest risk being arcing on the motors, followed by damage to the CCD from condensation. The risk of arcing in the motor mechanism was stated in the manual to be very low, below 0.001 Torr (1.33×10^{-3} mbar).

More likely, and worrying, was the risk of condensation of a very thin water layer on the CCD surface, which would deflect or absorb incident particles, or damage the detector, even if the motors were not in use. The exact pressure was not critical for

the propagation of electrons, as long as the chance of interaction with air particles was statistically low, and this was calculated to be negligible at the pressures and particle energies discussed here. For an electron of energy of around 77 keV, the range in dry air at 1.33×10^{-3} mbar is over 60 m, based on ESTAR range data.

With these points in mind, the aim for the experiments was to work at pressures below 1×10^{-3} mbar as an absolute maximum pressure and ideally at least an order of magnitude below this.

Since high vacuum conditions were not available at the time of the experiments, a turbopump was connected to the chamber which was capable of producing a vacuum level usually in the range of 10^{-5} mbar to 10^{-3} mbar depending on, among other things, the equipment inside the chamber, the length of time available to pump down and the temperature. This was judged to be sufficient for these tests.

5.2 Production

In order to simplify as much of the experimental setup production as possible, parts from previous experimental setups at MSSL were used wherever possible. This put constraints on the instrument that could be tested, and the tests that could be run. However, all the key parts were found, except the mask, which was manufactured at the University of Wolverhampton, and the mounting structures and some of the harnessing, which were manufactured in-house at MSSL. Each of the components and the reasons for its choice are described here.

5.2.1 Experiment components

Since a detector is a component that must be acquired off-the-shelf, it was necessary to first choose the most suitable detector, and then to design the experiment and its geometry around the detector available.

Detector

To demonstrate the mask concept, almost any position sensitive detector could be used. In fact, several aims of the instrument might be met by any detector (or an array of detectors) capable of simply collecting energy deposited per pixel over an arbitrary integration time.

However, more advanced CMOS detectors would offer more opportunity for individual particle detection, identification and angular estimation. Future experiments could, and should, utilise such a detector. Although the development that has gone into these more advanced detectors may give them more capabilities, it should be borne in mind that to extract and recover this information may require either or both of more processing power and time or a higher data rate via telemetry.

For these experiments, a spare high-end CCD was used: a back-illuminated, ion-implanted CCD64 produced by e2v, originally developed for the Solar X-ray Imager (SXI) instrument on the Geostationary Operational Environmental Satellite (GOES) spacecraft. This is an X-ray imaging detector with a very thin entrance window which allows it to detect electrons even at low energies; Bedington et al. [2012] described measurements of electrons down to approximately 500 eV using another CCD in the same series.

A back-illuminated CCD is initially manufactured in a similar way to a standard CCD, but is then processed to avoid the loss of sensitivity associated with the electrodes covering parts of the upper surface. The back of the CCD bulk is ground down so the depleted region is exposed and the back of the CCD becomes the new front, while the readout electrodes can be positioned behind the sensitive volume. This leaves an exceptionally thin entrance window, comprised only of whatever oxidised layer is formed on the new surface; this is therefore sensitive to a wider range of ionising particles than a standard CCD which would be very poor at detecting either X-rays or electrons. Figure 5.3 shows a diagram of the concept.

In a back-illuminated CCD, the exposed surface grows a thin silicon dioxide layer. At the interface between this and the bulk silicon, a potential well for electrons forms,

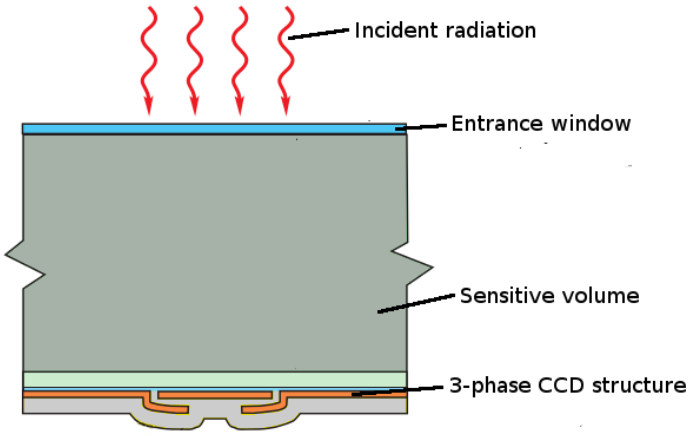


Figure 5.3: The back-illuminated sensor concept. The very thin entrance window provides very high sensitivity to low energy particles.

which makes low energy particles or photons which have a very short penetration depth difficult to detect. To increase the sensitivity to such radiation, the surface was treated with ion-implantation. When the surface of the semiconductor is exposed to higher-energy dopants by using a beam of ions from an accelerator, the depth at which the dopants are deposited and n^+- and p^+-layers are formed can be controlled. This is normally followed by an annealing step to reduce the radiation damage caused during the implantation. Ion-implanted detectors are relatively stable and can have very thin entrance windows.

The available CCDs were expected to be capable of detecting the particles of interest, based on previous tests with the detector at MSSL, but several relative weaknesses need to be borne in mind that limit the possibilities of the use of the detector both in the vacuum chamber experiments and in a possible flight instrument. Firstly, the CCD is highly sensitive to light. When it was run outside the vacuum chamber, the frames produced were saturated even when the CCD was covered. While this would cause significant problems for a flight instrument, in the vacuum chamber there were no light sources and this was not a concern for the proof-of-

concept. Secondly, the possible readout speed of a high-end CCD is limited. The listed performance of the CCD64 is given for an assumed readout frequency of 50 Hz or approximately 6 s for a full frame. Significantly faster clocking on the readout would be expected to increase the noise level. For the electronics used in these tests, the readout speed was fixed, and this informed the possible configurations of the test plans. In addition, the thermal noise levels and therefore the particle energy threshold of detection were of concern; according to the datasheet dark current varied with temperature proportional to $T^3 e^{-6400/T}$. For low-energy particle tests this limited the experiments to accumulating signal over an integration time, but since this was the intention of the instrument, this limitation was not a problem for these experiments.

The CCD64 series was characterised by Stern et al. [2004] using visible and X-ray sources, and to some degree by Bedington et al. [2012] for low-energy electrons. Based on the results from Bedington et al. [2012] and the stated optimisation of the surface for 200 eV photons, the dead layer was presumed to be close to 5 nm thick. This estimate is investigated further in section 5.5.5.

The CCD64 was supplied in a ceramic and metal package with three fixing lugs and a thermally conductive connection on the back to allow efficient cooling of the silicon itself. In addition, two AD590 temperature transducers were mounted in the package with very good thermal connections to the silicon substrate of the CCD which were used for temperature measurements. This package is shown in figure 5.4 with the shiny surface of the CCD in the centre, the bonded electrical connections above and below and the AD590s to the left and right of the detector.

This CCD64 is a full-frame 512×580 pixel CCD with an area of 8.2 mm×9.3 mm and 16 µm pixels. The 512×580 active region has 8 additional pixels at each end of each row to make the final frames 528×580 pixels. The pixels at the sides are not sensitive and are used to gain a value for the dark current of a frame. The CCD has a fill factor of 100% - there are no shielded areas on its surface where incident radiation is not detected.

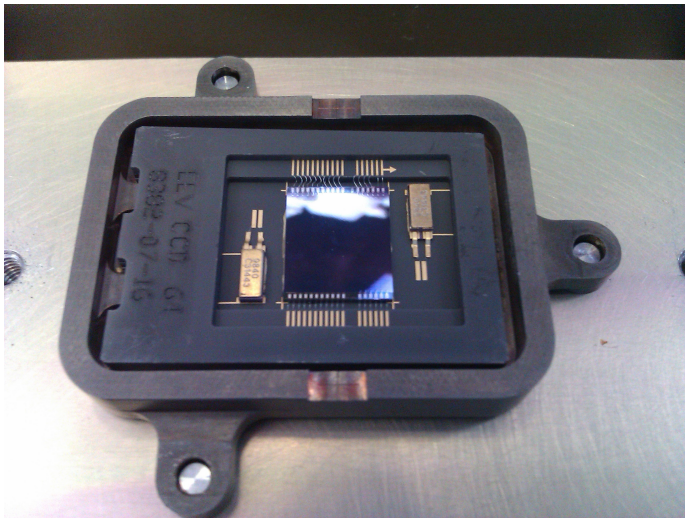


Figure 5.4: The CCD64 in its ceramic mounting. The CCD itself is in the centre of the image and the bonding wires for the clock signals and readout are shown above and below. The AD590 temperature sensors are shown to the left and right.

The electronics and software to produce signals to the CCD and provide readout capability to the device were reused from test electronics for the CCD47-20 from the Optical Monitoring Camera (OMC) instrument on INTEGRAL produced by Walton et al. [2003]. The electronics were split between those inside and outside the chamber: the amplification circuitry was positioned directly underneath the CCD which produced a signal reliable enough to be transmitted from inside to outside the chamber, where digitisation produces a 12-bit output. This was then stored by a Personal Computer (PC) running Interactive Data Language (IDL) 5.3.

Although several of this model of CCD were available, only one could be made to work reliably with the available electronics. This specimen did have a low level of suspected radiation damage and a number of column defects and bright pixels, but these were not judged to be insurmountable. The steps taken in the analysis software to deal with these anomalies are described in section 5.5.2.

Mask

The requirements for the mask were both an opaqueness to the particles to be used and mechanical stability in manufacture and use, bearing in mind the potential development into a flight instrument. Since the detector for eventual use was expected to be square, a square half-open pattern was chosen - a 16×16 PBA to be repeated 2×2 times on a scale proportional to the size of the detector.

The 32×32 size was also well-matched to the frame size of the CCD detector, the active area of which was, as stated above, 512×580 pixels. The area to be used for the reconstruction would be selected to be square, 512×512. This then equates to 32×32 pixels as the size of one grid element.

In addition, for this design, some support within the pattern is needed. The reconstruction properties of a URA-based mask are not affected by the proportion of each open element in the array which is actually open, beyond the geometric factor of the instrument and resulting particle count rate. Therefore, this support was effected by producing holes at a scale of one half the size of the array element and the open fraction of the mask itself was be one eighth of the aperture area. This meant that the holes were to be the size of 16×16 pixels, or 256 µm square and the hole aperture area was 16.4 mm square.

Initial attempts at producing a mask in-house at MSSL using a high-power laser cutter were unsatisfactory. The scripts were produced for several URA and PBA designs with different structure widths, but the laser cutter encountered difficulties in manufacturing them. The densest material available was copper sheet, and while it was hoped that the laser cutter might be used to produce such a fine structure, it was unable to handle sufficient metal thicknesses. Thinner copper foil was deformed by the laser and not usable, and thicker pieces would need multiple passes of the laser cutter at different focal lengths, in which time the cutter had often lost its alignment. Figure 5.5 shows the result of an attempt which was followed by the use of high pressure air to blow out the loose particles. Unfortunately, many of the pieces were not loose enough to be able to be freed without seriously damaging the

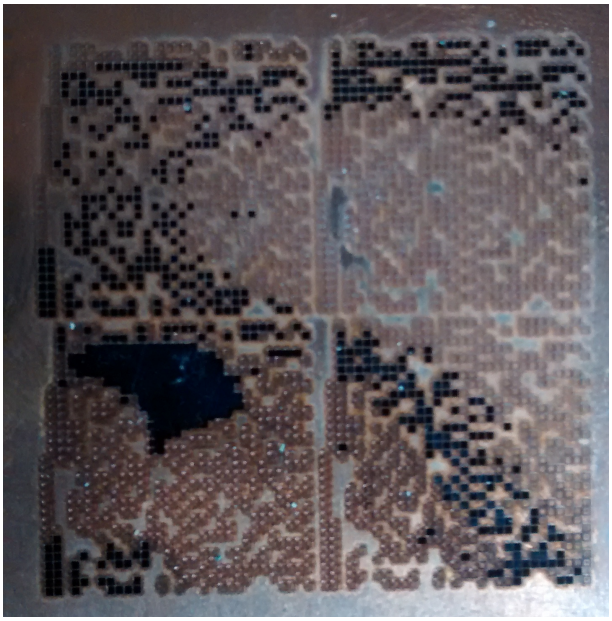


Figure 5.5: An attempt to make a PBA mask out of copper shim with narrow supporting structures between each hole, which was unsuccessful due to the limitations of the laser cutter at MSSL.

rest of the structure, even after two passes with the laser cutter. Therefore, external manufacturing options were investigated.

The thickness chosen for the mask was planned so that the ratio of hole depth to width would perform a degree of collimation on any surrounding particles. For a mask with 256 µm holes positioned 4 cm from the detector, a mask thickness of 400 µm is sufficient to filter out some of the incident radiation not aimed at the detector. This would not significantly affect the mask performance in the vacuum chamber tests, where the position of the source is known and controlled to be near the centre of the field of view. However in unknown environments this approach would cause complications with deconvolution as the shadow pattern changes more significantly when part of the source is off-axis.

Ideally, a dense material such as tungsten was planned for the mask, but due

to the complexity of machining tungsten or other tungsten manufacturing methods, lead times and costs for such a mask were prohibitively long. The prototype mask used in these experiments was manufactured by the University of Wolverhampton by laser sintering a mixture of copper and tungsten particles to form the mask pattern in a fully-dense pseudo-alloy several millimetres thick. Each mask was then sliced from the thicker block and was measured to be approximately 470 µm thick with a total mask area of a square of approximately 36 mm along each side.

A scanned image of one of the masks used is shown in figure 5.6, where the granular texture caused by the manufacturing process can be seen. This granularity is shown in figure 5.7, as viewed under a microscope. Since this would be very hard to model accurately, the mixture was approximated in simulations by a homogeneous mixture with perfectly square holes, but the irregularities in the hole shapes caused by the process were noted when examining the sharpness of the resulting images.

The manufacturing plan was for approximately a ratio of 1:3 tungsten to copper particles, which gave a planned density of approximately $11.3\,\mathrm{g\,cm^{-3}}$. However, each mask was measured as approximately 4.6 g or around $9\,\mathrm{g\,cm^{-3}}$, very similar to the density of copper. This could be because of the low proportion of tungsten used in the manufacturing process, or the rough texture of the surface may have made the volume measurements inaccurate.

Since the material was not as dense as planned for a final instrument, a mask thickness of approximately 0.94 mm could be created by clamping two masks together in front of the detector, although the necessity for this would depend on the incident particle spectrum. Clearly, the thicker the mask needs to be, the smaller the angles of incident particle that could be accepted and analysed. Additionally, the imperfections from the manufacturing process shown in figure 5.7 would reduce the effective hole size slightly further.

For the particle energies used in these experiments, only one mask was required. The highest energy electrons produced by the carbon-14 source that was used would be absorbed completely by 60 µm of copper.

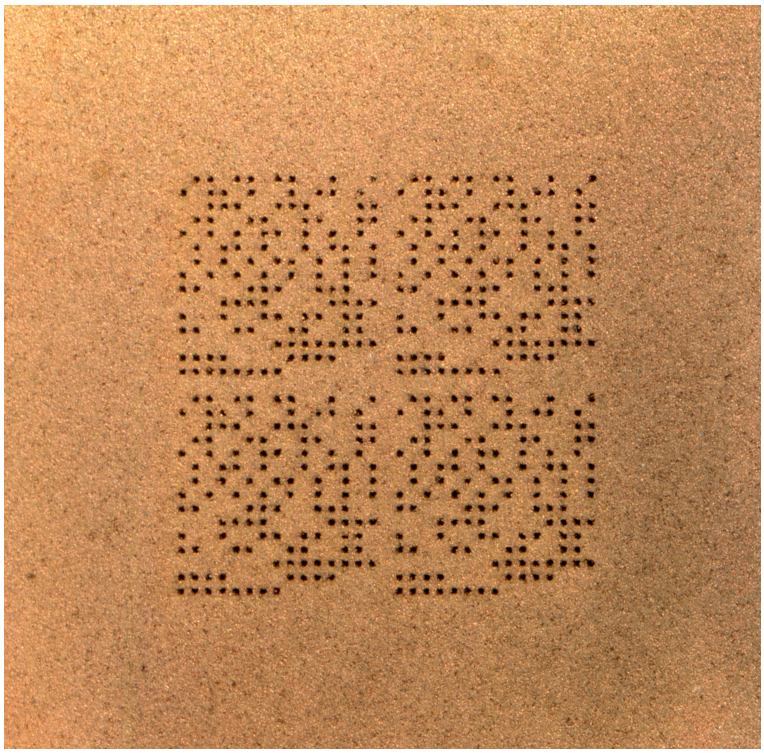


Figure 5.6: A scanned image of one of the masks used

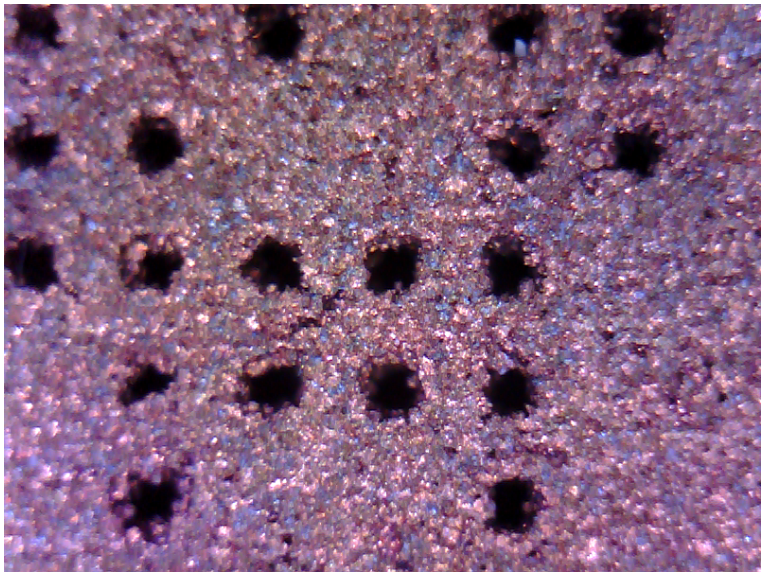


Figure 5.7: A magnified view of part of one of the masks used

Mounting for the CCD and mask assembly

Mounting for the CCD needed to be modified and the mask mounting added. This part of the setup was designed and manufactured by the workshops at MSSL and is shown in figure 5.8 with the mask fixed 4 cm from the surface of the detector.

To avoid damage to the detector and sensitive electronics in an experiment or flight configuration, the shielding needed to cover the silicon sensor and the readout electronics at least in the parts handling the signal before it is amplified. A mechanical design that could shield a small area of circuit board and components would make the concept of a simple instrument that can be self-contained and added to almost any spacecraft far more feasible. It must however be borne in mind that the shielding will likely be a major fraction of the mass of a flight instrument and should be minimised geometrically as far as possible.

In the chamber setup, the mounting from a previous assembly was adapted in such a way that the plane of the CCD was vertical and the amplifier electronics was

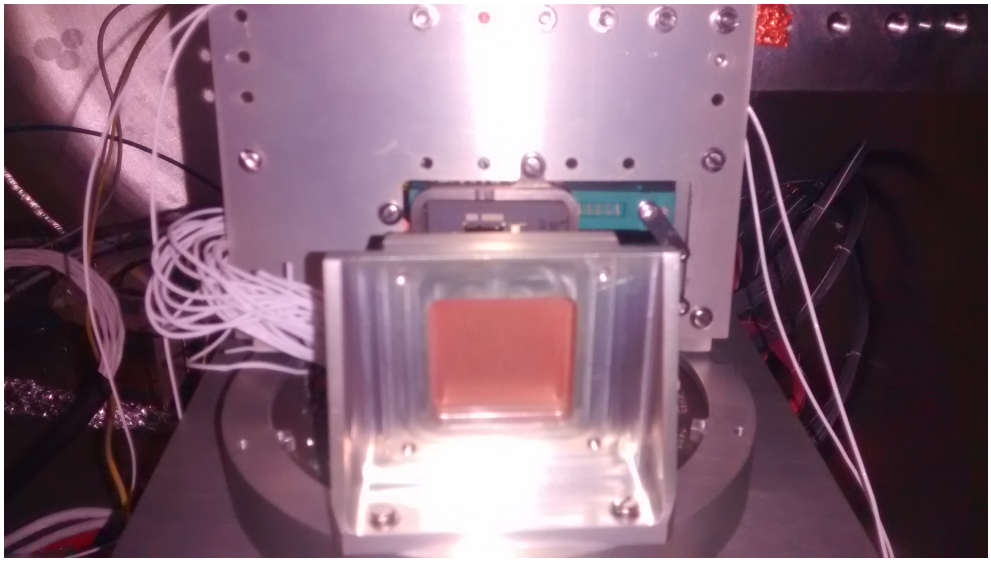


Figure 5.8: The mounting arrangement of the mask, 4 cm in front of the CCD. The circular Printed Circuit Board (PCB) with the vacuum side electronics can be seen below these mountings and the cold finger described in section 5.2.1 is visible at the top right of the image.

positioned horizontally, where they would be subject to only a low radiation dose.

Shutter mechanism

Initially, the use of a shutter was considered, but this was neither available nor could be acquired, so one of the major experiment constraints was dealing with the restricted readout speed. The CCD and software available would permit a readout time of around half a second to a second. During this time, the CCD would continue to detect particles and accumulate dark current which would result in a deformed image which could not be accurately deconvolved. This deformation could be minimised by increasing the ratio between exposure and readout time. However, long exposures of the ten or more seconds required resulted in an increase in the thermal noise accumulated and therefore significantly increased the cooling requirements necessary to get a good image, and prevented identification of individual particles at the energies that were used for longer exposures.

In the experiments described here, measurements were variously taken at 10 s, 100 s and 1000 s exposures. This exposure time was necessary not only to reduce this effect, but also to gather enough particles for the flux rates from the radioactive sources.

The shutter would also have been used as a way of protecting the detector from radiation outside exposure times. This would be used in the case of energetic or high flux particle sources that could cause damage to the CCD. This was not required in the final design, since the particle sources used were low energy and would not be a danger to the detector.

Thermal control

Since the CCD64 was only produced in a small volume for a specific application, the datasheet provided gives few specifications for the optimal operational temperature. However, the readout noise is defined for temperatures from 140 K to 235 K. In addition, the datasheet for a similar CCD, such as a CCD42, gives a minimum

operating temperature of 153 K and additionally specifies that cooling or heating rates should not exceed a rate of 5 K/minute. Within these ranges it is desirable to reduce the temperature as far as possible to reduce background noise in the images.

Thermal control was possible using a liquid nitrogen-cooled aluminium finger connected to the backplate of the CCD by a clamped piece of copper braid. In addition, two resistive heaters were attached to the backplate, although in practice these were not used beyond testing them. It was judged preferable to obtain measurements at the lowest possible temperatures (in some cases down to 188 K) rather than favouring a constant temperature, and it was clear once the experiments were underway that there was no risk of the CCD being cooled at dangerously fast rates.

The back of the mount used for the CCD is shown in figure 5.9. The CCD in its housing was screwed onto the other side of this plate and two zero insertion force (ZIF) connectors were used on the rows of pins at the back of the CCD to connect the sensor to the electronics under the mount. The copper piece shown was connected to the back of the CCD package where the housing provided a thermal link to the CCD itself. Also shown along the top of the plate are several screw holes which were used to tightly attach the nitrogen-cooled finger via a piece of copper braid, to try to cool the CCD as far as possible, once the experiment was under vacuum. This mounting plate was insulated from the electronics plate to which it was attached by the use of Kapton washers to avoid excessively cold temperatures reaching the amplifier board electronics.

Although the existing control electronics and software were not previously connected to the two AD590s on the CCD, connecting them was relatively simple. A simple resistor and voltmeter setup, shown in figure 5.10, was sufficient to get accurate thermal readings of the CCD sensor itself.

The temperature reading was directly visible on a voltmeter, at a scale of $10\,\mathrm{mV\,K^{-1}}$. When the CCD was running, it was possible to observe the 0.1 K rise and fall as the readout and flush sequences ran. During the experiments, one of the sensors appeared to become unreliable, so the other was used as the temperature monitor

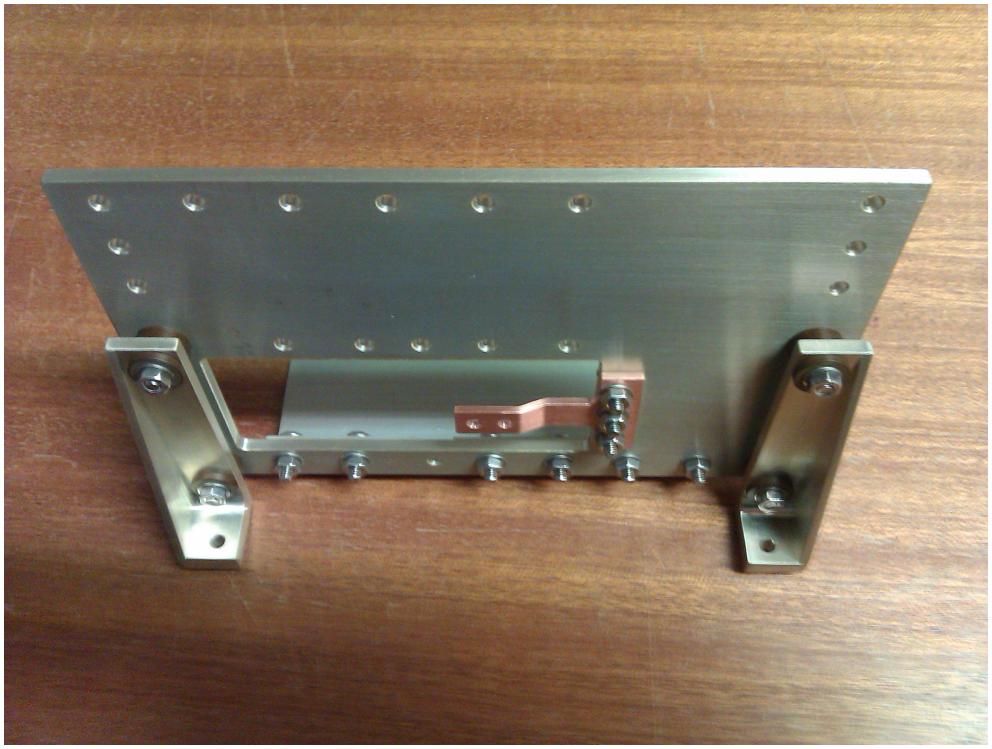


Figure 5.9: The back of the mounting piece to which the CCD was attached showing the mountings for thermal components of the experiment. The CCD was positioned in the hole with its surface facing away from the camera and the copper thermal connection mounted to its back. The nitrogen-cooled cold finger was bolted to the row of holes along the top with a section of copper braid sandwiched in between to increase the thermal connection.

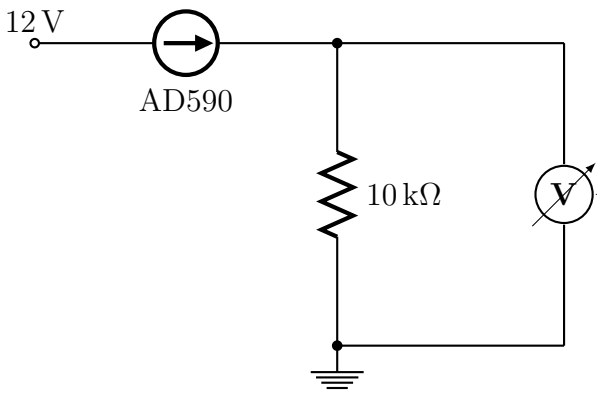


Figure 5.10: Electrical schematic of the AD590 CCD temperature readout. The voltage was read from the screen of a Fluke multimeter and converted to temperature.

for the rest of the measurements without any further problems.

Motors for alignment

Extended sources were possible to simulate by moving the source on a mounting table which was positioned in two directions by Burleigh Inchworm motors. These are piezoelectric controlled linear motors capable of greater than 10 nm accuracy - far in excess of what was necessary. The motors were controlled with a PC running software to control the high voltage (HV) supply in the Inchworm motor controller via an Industry Standard Architecture (ISA) PCB.

5.2.2 Particle sources

While a source at a finite distance (of the order of the size of the vacuum chamber) from the detector would not provide an angular distribution of the incident particles, the plan was that a source placed at a set distance from the detector could be convolved once the mask had been scaled proportionally to the image produced on the detector. In this case, the deconvolution will produce an image of the source in

position space over a plane at the measured distance to the source, rather than a plot in angular space.

Three sources were used in the setup: two radioactive beta sources, and a modified MCP-based low-energy electron source.

The radioactive sources

The radioactive sources available at MSSL were relatively low activity which meant that individual particles would need to be identified. Therefore the particles needed to be of an energy that would produce a signal in the CCD pixels which would stand out above the noise in the image frames. A mounting was designed and produced by the workshops at MSSL to hold the radioactive sources in place on the motor table.

It was hoped that a strontium-90 source with a Q value of 546 keV which decays to yttrium-90 with a short half life of 64 hours and a Q value of 2.28 MeV could be used, so producing two higher-energy electron spectra, but calculations indicated that the source available would have a very low particle flux that would make deconvolution difficult. The most recent measurements taken of the source indicated a current activity of 353 kBq - over three orders of magnitude less than other available sources.

The initial experiments were run with a samarium-151 beta source, which has a Q value of 77 keV, with various maskings in front to modify the spatial distribution. Although the geometry of the source itself was known to be 14 mm in diameter, a thin cover of unknown thickness appeared to be used over the front to contain the source. The effect this had on the energy spectrum and particle flux was unknown, although a very limited estimate of the energy spectrum was attempted from the CCD data. However, the flux rate of measurable hits at the CCD temperatures used was lower than expected and a higher energy source was planned to attempt to accumulate a larger number of particle hit data above the background threshold.

Further experiments were run with a carbon-14 source with a Q value of 156 keV.

In this case, the particle numbers measured were still not enough for deconvolution within the time constraints of the nitrogen-cooling. The electronics available permitted a maximum of 50% of the time to be used for exposure, and a nitrogen dewar could maintain sufficient cooling for approximately 10 hours before the pressure dropped too low and the temperature rose again. In most cases it was at least two more days before the dewar could be refilled and reused, so the exposure time for a set of tests was usually limited to 5 or 10 hours.

The MCP-based electron source

The count rates of observable particle hits produced by the radioactive source setup were lower than expected, so in order to provide a clearer demonstration, a further series of tests was planned using a modification of an already-existing MCP-based electron gun setup. The benefits of such a setup for the purposes of clarifying the behaviour of the instrument would be clearer mask patterns, because of a more predictable and parallel beam and because of far higher flux rates.

This was constructed at MSSL using four ultraviolet (UV) Light Emitting Diodes (LEDs) behind an MCP to create photoelectrons and an avalanche of further electrons. The MCP requires a potential difference of around 2 keV to work and a further electrode grid in front at 0 V is needed to accelerate the generated electrons to the required energy from the negative voltage at the front of the MCP.

The setup in place already at MSSL was composed of five selectable Zener diodes rated to 50 V each to push the voltage at the side of the MCP from which the electrons emerge to a value from 0 V to −250 V, which accelerated electrons towards the grounded grid at corresponding energies up to around 250 eV. Since higher energy values were necessary for the CCD detector to be able to detect them because of the dead layer on the surface, it was decided to increase the energy range of electrons as far as possible using the hardware available.

Modifications to allow these changes and to move the source into the vacuum chamber containing the coded aperture setup were relatively minor. To increase

the maximum possible electron energy to above 2 keV, both the HV harnessing and the Zener diode switchbox needed to be adapted. The upper limit of energies that could be reached was dependent on the available parts for the HV harnessing at the available time - the most suitable were Safe High Voltage (SHV) connectors rated to 5000 V. A maximum of −4000 V was used in the chamber, both for safety and because of the values of the Zener diodes that were available to modify the switchbox.

The switchbox circuitry was modified by replacing each 50 V Zener diode with two 200 V Zener diodes to produce the circuit given in figure 5.11.

In addition, mechanical considerations needed to be taken into account. The MCP as used for previous testing was bolted to a metal plate. Once the MCP was positioned in the chamber and aligned by eye, it remained in the same place throughout the tests. The distance between the MCP front grid and the CCD was measured as 21.1 cm using a hand-held laser rangefinder. The MCP on its mounting relative to the mask and CCD is shown in figure 5.12.

To operate this setup the necessary number of Zener diodes must first be selected and then the correct HV supply turned on. Some experimentation was necessary to find the voltage limits of the MCP in this setup to avoid reaching a current limit of 100 µA across the MCP.

In most cases, the MCP was run with a voltage of 2000 V across it, and the potential difference between the MCP face and the grounded grid was varied in 400 V steps, which were intended to produce equal numbers of electrons of 400 V, 800 V, 1200 V, 1600 V and 2000 V. However, when the MCP experiments were run with the mask removed, the detector would saturate when using these voltages, and the flux was lowered by setting approximately 1800 V across the MCP.

CCD data were taken using the MCP source both with and without the mask in place. The inside of the chamber with the mask removed is shown in figure 5.13.

The grounded grid at the front of the MCP source was a diamond-shaped pattern, with wires across at two angles, with a total area of 75 mm×39 mm. The diamonds

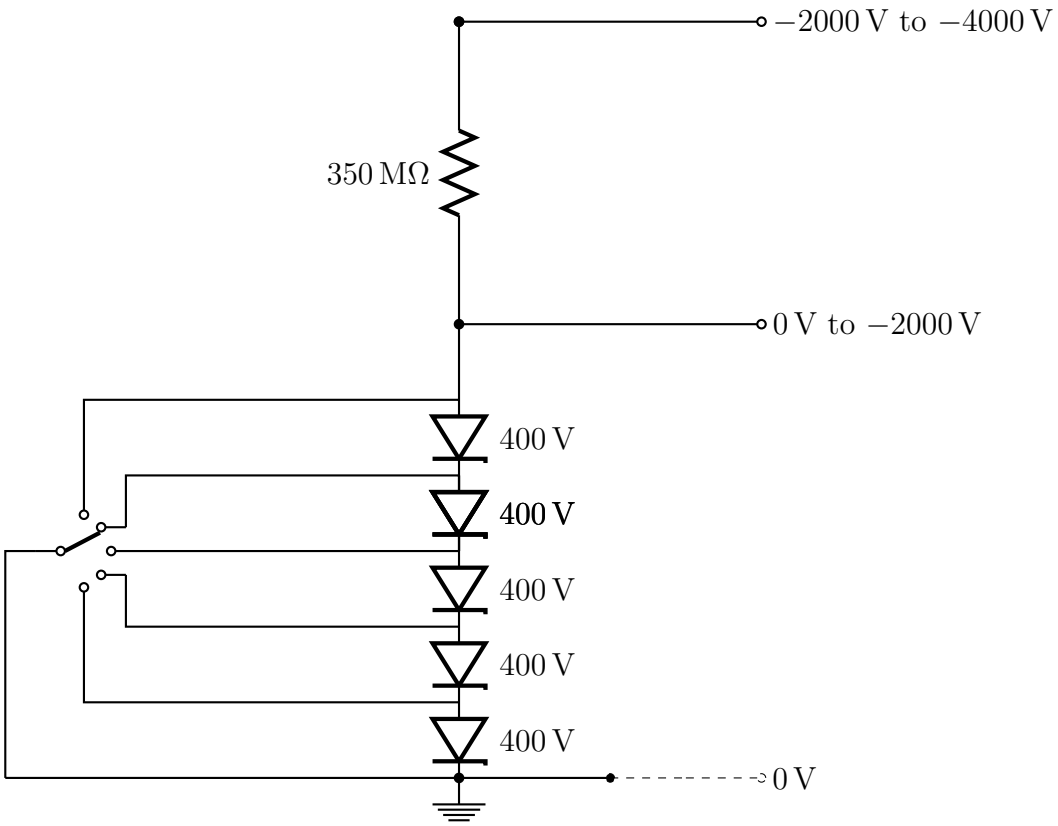


Figure 5.11: Electrical schematic of the electron gun energy selection where the MCP in operation is represented by a 350 MΩ resistor. The MCP and grounded grid were both fixed within the vacuum chamber, and the HV supply, switchbox and grounding were outside the chamber.

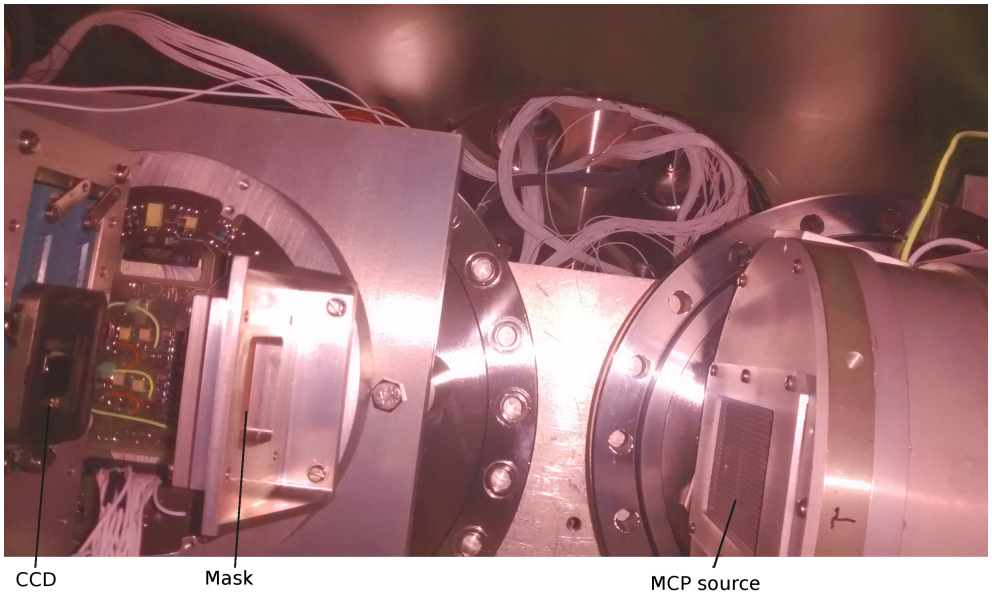


Figure 5.12: The MCP experiment setup in the chamber. The setup is viewed from above and the door to the chamber is off the right hand side of the image.

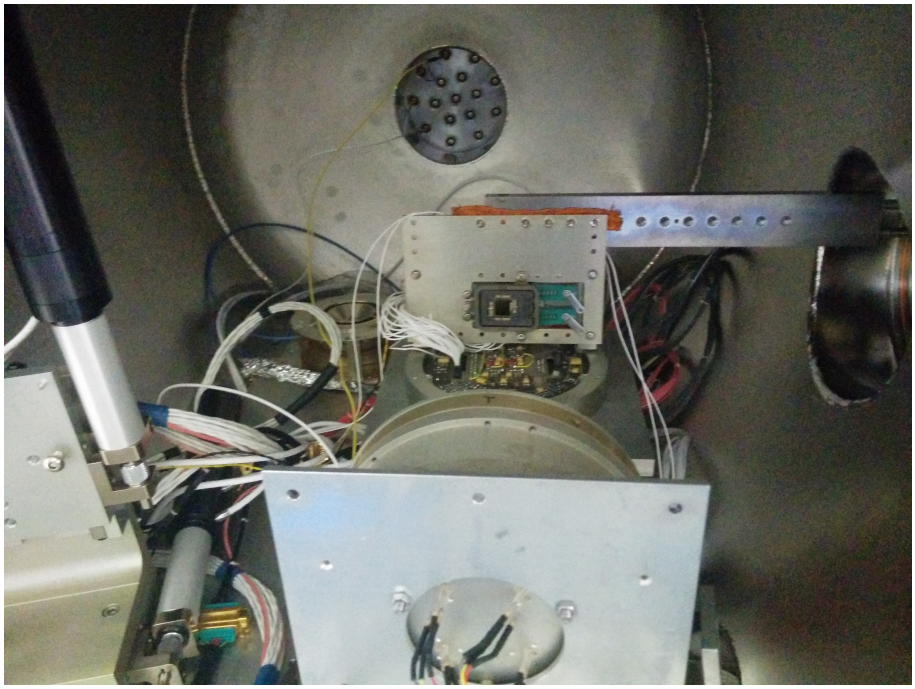


Figure 5.13: The MCP experiment setup with the mask removed in the chamber

were measured as approximately 3.25 mm×1.85 mm, with the angles of the grid lines at approximately 29.6° from the horizontal.

The MCP electron source had been partially characterised previously in its lower-energy configuration. A CEM-based detector was scanned over the beam. In addition, an attempt was made using a top hat instrument to characterise the energy and angular distribution of the produced electrons. Unfortunately, these data are of only limited use, since the top hat itself lacked full characterisation.

Once the higher-voltage, higher-energy setup was produced, the forward element of the electrons' momentum was expected to increase, and the beam to become more parallel. The angular and position distribution of electrons from the source were not well known before the testing, and further characterisation was not possible within the time the facilities were available.

It was hoped that the beam of electrons would be close enough to parallel that the mask and detector could be operated in the 'correct' configuration - measuring angular distribution rather than position distribution, and with no need to scale the mask shape for deconvolution. This could then be used to identify the angular size and distribution of the source. For this to work, the electrons from the edge of the MCP must still illuminate the whole of the detector. In this case, if the MCP was a well-aligned and evenly-distributed source, the electrons would have needed to be within 5° of perpendicular to the surface. The lack of knowledge about the source became more significant when it became apparent that the position and angular distribution was significantly more complex than expected. To gain clear knowledge about a source, at least one of the shape and the angular distribution must be known. Since both angular and position distribution were not known, recovering the source through deconvolution could not be done accurately.

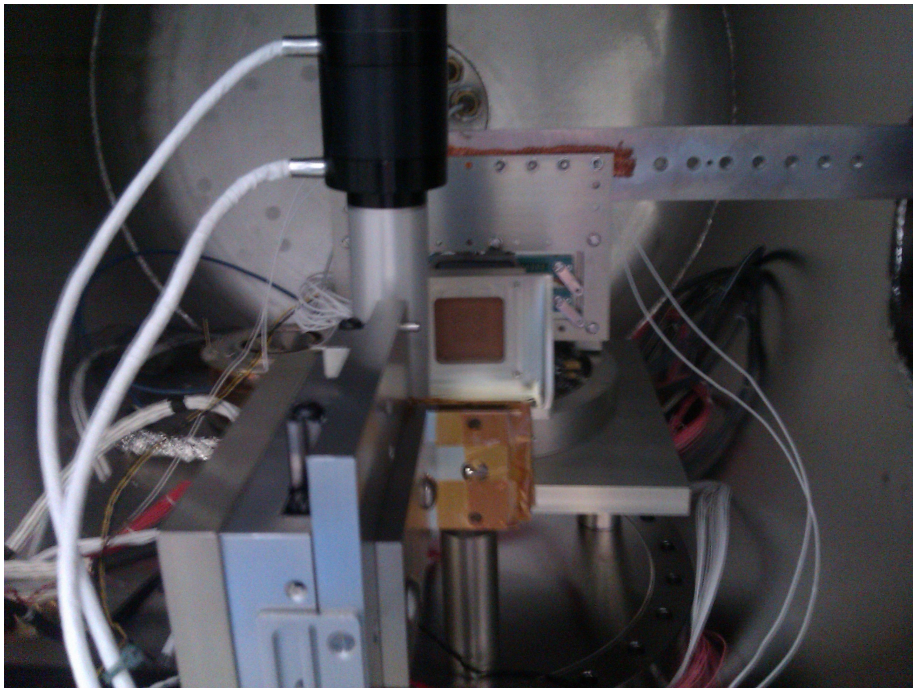


Figure 5.14: The vacuum chamber setup

5.3 Method and control

A photograph of the final setup in the chamber is shown in 5.14, with the motor and source in the foreground and the mask covering the CCD behind it.

Outside the chamber, there were two input and output channels to and from the measurement part of the system when the chamber was pumped down: the thermal control and the CCD control.

The cold finger was attached to the outlet on a large dewar of liquid nitrogen stored outside the laboratory and the nitrogen was allowed to flow through the finger and out again. After the first few runs it was determined that one of the three available dewars was significantly more efficient for this purpose than the others, and this one was used in preference for the rest of the experiments. With this setup,

it was possible to work below −40 °C for a stretch of approximately 8 hours.

The CCD signal generation electronics were controlled from the IDL command line. The exposure time was set for a continuous series of measurements. Although the readout time was limited to under a second, the time between exposures was the same as the exposure time, due to the functioning of the electronics. This exposure time was variously set to 10 s, 100 s and 1000 s. Longer exposures than 1000 s were not possible with the available electronics.

In addition, the source could be controlled: the position, for all tests, and the LED and HV power supplies for the MCP source.

5.3.1 Data collection

Data from the CCD was collected by the IDL 5.3 software used for CCD control, in the form of a 528 by 580 array. Each 12-bit digital number was stored as a two byte value and exported to a separate data file for each frame, with header values such as date, temperature and exposure time recorded in the filename. These data files were read into and analysed using Matlab, usually after concatenating a whole data collection run into one data file to produce a datablock in Matlab of all the relevant frames in layers.

5.3.2 Power supply and harnessing

The CCD electronics to be used were already prepared with the required harnessing between the detector package and the amplifier electronics, and between the clock generation electronics and the PC. However, the connection of the electronics inside and outside the vacuum chamber was needed; the harnessing used previously contained a Douglas feedthrough embedded in another flange which could not be removed from another vacuum chamber.

The vacuum chamber available for the experiments was only compatible with certain flanges. One of these was fitted with a number of 37-pin circular Deutsch

connectors. Connectors to fit to these which were found to be compatible with the ground test harnesses from the PEACE instrument from the Cluster mission which was launched in 2000. The harnessing was designed to be as quick and cheap to manufacture as possible to fit with these harnesses.

One downside of the use of these cables was their length: the vacuum side cables were 1 m long and the air side cables were 2 m long. However, they contained several twisted pairs of wires, some of which were shielded and could be used for the more sensitive signals.

A set of interconnected cables through three Deutsch connectors on the feedthrough was selected as being sufficient for the purposes of the connections needed. On each side of the feedthrough, these led to a fully-populated 50-pin D-subminiature, a 25-pin D-subminiature with 16 pins populated, a 15-pin D-subminiature with 10 pins populated and a fully-populated 9-pin D-subminiature. It was not possible to modify these harnesses, so the existing wiring needed to interface with them.

The harnessing to the CCD amplifier board electronics and the control electronics needed two high-density D-subminiature connectors, of 26 and 44 pins. A suitable pair of harnesses was found and modified to connect the 44-pin D-subminiatures with the 50-pin connectors, but a harness needed to be manufactured to connect the 26-way connector. It both needed to be connected through the available 25 and 15 pin connectors and to avoid distortion of the sensitive analogue signals. For example, in the design of the harness, the signal pins were matched to wires that were in shielded twisted pairs in the Cluster harnesses.

These new harnesses were manufactured by the electronics workshop at MSSL.

In addition to these, the heaters and AD590s were separately wired to the available 9-pin D-subminiature connector from the Cluster harness, using two pins (out and return) for each of the two heaters, and two pins (out and return) for each of the AD590 sensors. The AD590 connections were used with a resistor and voltmeter as described in section 5.2.1.

5.4 Experimental sequence

A series of tests with different electron sources was run. The results are described in section 5.5.

Initially, a samarium-151 source was used, but when the flux rate appeared to be insufficient, a carbon-14 source was tried instead. Since this did not produce sufficient flux either, once several measurements had been taken from each, another lower-energy electron source was proposed. This had been recently produced at MSSL from UV LEDs and an MCP. Some modifications were necessary to increase the energy range, and a series of measurements were taken using this source, with and without the coded mask in place. This section describes the order in which measurements were taken in the vacuum chamber.

5.4.1 Initial samarium-151 tests

Initially, the setup was run in vacuum without an electron source, to produce background frames at a range of temperatures. Although this was useful in understanding and checking the system, the frames produced were of limited use for the rest of the experiments. The gain on the CCD64 was variable between frames at the same temperature, and the temperamental nature of the synchronisation of the CCD signals was found.

Once this set of test had been run, the samarium-151 source was put into place with a pinhole directly in front of it. The flux was found to be surprisingly low. Since there were data from the earlier tests without the source in place, it was established that the few bright spots on these frames were signal rather than background, for example from cosmic ray effects. In an attempt to gain enough signal to attempt a reconstruction, these tests were run repeatedly over a number of dates, until it became apparent that the number flux of observable particle was too low to produce a useful image.

5.4.2 Carbon-14 tests

A possible cause of the low number of measured counts was the energy spectrum of the samarium source. To understand the interaction between higher-energy electrons and the CCD64, another radioactive source with a higher Q-value was chosen: carbon-14.

Initially this was tried from a distance of 16 cm with the pinhole in place, but this also produced a unusably low number of identifiable hits. In order to increase the number of hits, the pinhole was removed and the full disc was used.

In order to quantify the low flux, the source was positioned approximately 1 cm away from the CCD, with the mask removed. This was close enough that the CCD would be expected to be approximately evenly illuminated and the flux per area of the source would correspond to approximately the same flux per area reaching the CCD.

5.4.3 Further samarium-151 tests

Since neither radioactive source produced the flux that was hoped for, the data collection for these tests was completed by taking measurements for the samarium source to match those taken from the carbon source. The samarium source was placed about 1 cm away from detector to produce an equivalent frame for count rate and energy spectrum estimation as for the carbon case.

The mask was replaced, and the samarium source moved back to the mounting on the motors. Since the pinhole aperture in front of the source had not produced close to the flux required, a more distributed source pattern was attempted, with most of the disc exposed, and a copper wire across it to add some further structure. This was tried for three days of tests to see if the summed flux could be used to reconstruct this distribution, but this was insufficient.

5.4.4 MCP source tests

The low flux caused serious difficulties with the radioactive sources available. Using the MCP electron source produced at MSSL was proposed. This was, in essence, an MCP positioned in front of a set of UV LEDs. When the LEDs were turned on, and a sufficient voltage was put across the MCP, the UV illumination would create photoelectrons at the back surface of the MCP and the gain of the MCP (adjustable by voltage) would create a high flux of electrons. These were then accelerated through a further electric field to produce the energy needed. This assembly is described in further detail in section 5.2.2.

In its modified form, the MCP source was, in theory, capable of producing electrons of 400 eV, 800 eV, 1200 eV, 1600 eV and 2000 eV. The number flux could be adjusted by changing the voltage across the MCP and therefore its gain. The angular distribution of the electrons produced was unknown, but expected to be near-parallel and dependent on the electron energy. Adjustments in the position of the source could not be made without breaking the vacuum, since the MCP source and its mounting were too large to fit on a motor platform.

Measurements were taken with the mask in place for many combinations of LED voltage, MCP voltage, and acceleration voltage.

Finally, in order to produce flat field frames at each electron energy, a test run was made with the same alignment, but without the mask in place. The voltage across the MCP was lower than in the previous tests, since otherwise the flux was high enough to saturate the CCD.

5.4.5 Other tests

An attempt to use light was made after the tests, using an LED. This would demonstrate the concept of the mask producing a shadow pattern, but this would ideally have needed to be done in the vacuum chamber.

This presented several difficulties - producing and running a vacuum compatible

setup was judged to be an inefficient use of resources, chamber time and nitrogen. Cooling the CCD outside the chamber would have risked causing damage to the detector, since condensation would form on and around it, as well as likely to form on other connected electronics. Forming a completely dark box caused further difficulties, and recovering a usable image was unsuccessful. Although this would have demonstrated the ability of the coded aperture technique, its use with photons over a wide range of energies has been demonstrated previously both in space and on Earth.

Further adjustments of the position and orientation of the MCP source were not possible when it was under vacuum, and opening the chamber sacrificed measurement time, so it remained in one position for each of the tests.

5.5 Results and analysis

The data were of mixed quality. The gain of the CCD electronics was highly variable and the synchronisation of the electronics was sometimes lost. This presented challenges in the image processing and to some degree limited what was possible with the data. However, in all experiments, results were gained which gave useful information about the properties of the sources that were used.

5.5.1 Expected results

Before the tests were run, some estimates were made of the predicted outcomes based on what was known of the sources. In the case of the radioactive sources, the activity levels had been recorded; in the case of the MCP-based source less was known, since it was newly-adapted, but some estimates of particle number fluxes had been made.

Radioactive sources

The two radioactive sources, samarium-151 and carbon-14, were available at MSSL and their activity had been measured when they were acquired.

The samarium-151 source at MSSL was measured as having an activity of 925 MBq in 1987, and samarium-151 has a half life of 96.6 years. Therefore, in 2014, the source could be supposed to have an activity level of 762 MBq. If the detector is 20 cm from the source and has an area of 76 mm^2, it will receive 0.015% of this activity. If the coded mask is in place, this fraction will be divided by 8. Therefore, the CCD could expect a hit rate of 14 450 s^{-1}. Since the number of active pixels is 296960, approximately 0.049 /s/pixel or, for a 100 s exposure, a total of 4.8 hits/pixel would be expected. This value would be reduced when any further masking or shielding is placed in front of the source.

Similarly, the carbon-14 source was calculated to have an activity of approximately 100 MBq, which would give 0.63 hits/pixel in a 100 s exposure. Although this is significantly fewer than the samarium source, the beta particle energies are higher, and more of them would be seen against the thermal and electronic noise in the CCD frames. This noise had a standard deviation of around 40 digital numbers in the cold (below −40 °C) frames.

MCP source

The LED and MCP setup was supposed to produce a number of particles which depended on the voltage across the MCP.

As an example, if the MCP were set up to produce 2 million collimated electrons per second over an area of 7 cm by 4 cm in the direction of the CCD, 54.3 thousand of them would hit the detector each second. This is equivalent to an average of 0.18 hits/pixel/second, or 0.023 hits/pixel/second with the mask in place. If 2 keV electrons deposit all their energy, they would produce about 600 generated electrons each which with an electronics gain of 25 is about 24 Analogue to Digital Units (ADU). In 100 s, this would produce average of about 54 ADU/pixel. This estimate is high

enough to expect a detectable signal and low enough to avoid the necessity of shielding the detector in any way. If the signal were either too weak or too strong to measure, the voltage across the MCP could be adjusted, limited by the voltage that would cause a current that might damage the MCP.

The expected detectable electron energies were expected to start at around 500 eV, as measured by Bedington et al. [2012]. The Zener diode settings of 800 eV to 2000 eV were expected to give a clear signal.

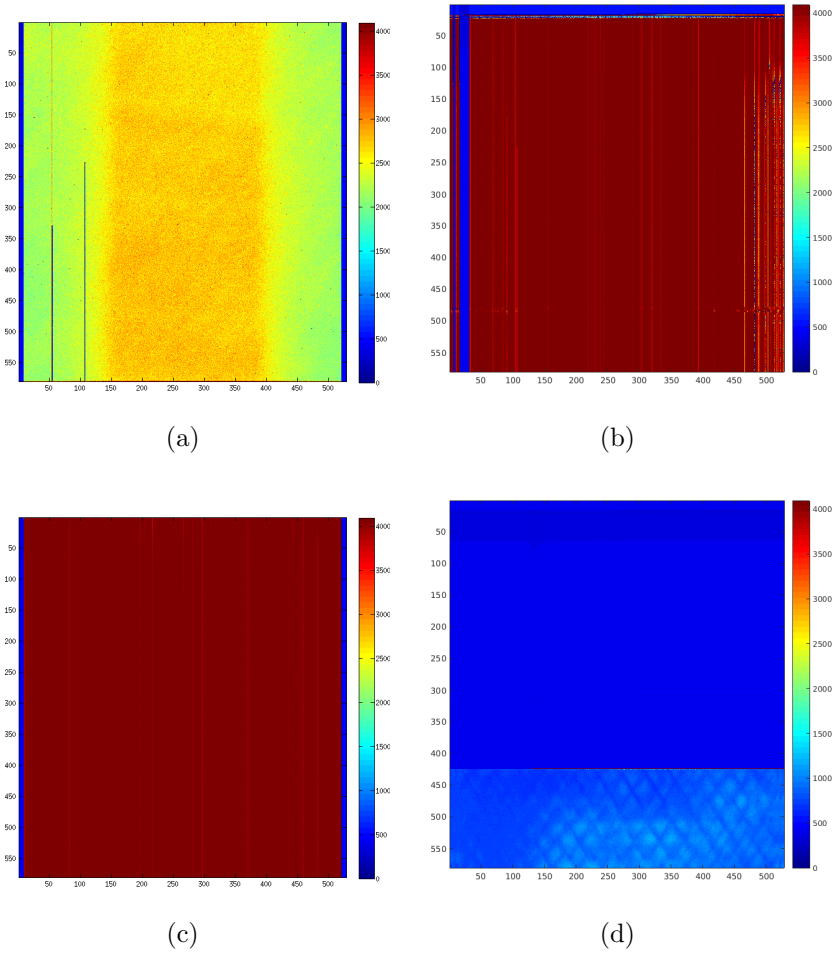
5.5.2 Data processing procedures

The raw CCD frames shown in this section are unprocessed, except for scaling the colours in some cases to highlight the details. However, for further processing and analysis, the damaged parts of the frame and the frame background level and noise needed to be accounted for. The data processing was modified slightly for each frame or series of frames as required, but the steps that were used for some or all of the data frames included the following ones.

Damaged frame removal

A few frames were damaged, especially at higher temperatures, within each run. This was assumed to be a problem either with the damage on the detector, or with the CCD digital signal synchronisation. The frames were simply discarded without further investigation; they were identified either because the average pixel value is far from what is expected or by eye. In addition, those frames taken at higher temperatures within a run had higher background noise and were generally lower quality. These were identified by plotting the mean pixel value per frame and removing frames above a certain threshold, depending on the noise requirements of the radioactive source and the investigation.

Figure 5.15 shows some examples of frames that were clearly damaged and unusable and were identified by eye and removed.



(a) (b)

(c) (d)

Figure 5.15: Four examples of damaged frames that were removed after inspection.

Damaged pixel and column removal

The CCD used for all tests was slightly damaged, with several bright pixels, a scratch or other mark and one and a half bright columns and some suspected light radiation damage on the middle third of columns and more severe damage at the upper edge. These bright pixels were removed from the frame and replaced by the local average where they were clearly incapable of providing a useful signal.

The radiation damage in the centre third of the columns was mild and for constant exposure times appeared to be acceptably removed by subtracting the average background at similar temperatures. The more severe damage on the top rows of the image was more difficult to remove from the frames entirely, but since the exposure area used for reconstruction should be the same shape as the mask pattern, in this case a square PBA, it was preferable to use only a 512×512 pixel area of the data and the top 62 rows could be removed, reducing the problem significantly.

Averaging

For long runs with very few particle hits, the stack of processed frames was averaged to make an estimate of the expected value of each pixel, and a scaled value of this averaged frame was subtracted from each exposure. This took into account the different behaviour of each individual pixel. In the case of runs which had very variable gains, the average of the individual frame made the most suitable flat-field background since the behaviour of each pixel between frames was erratic.

Identification of particle hits

For the radioactive sources, where particles could be identified individually above the background noise, the particle hits were at first identified by eye as either a point, a blob or a short track.

When the sources were moved closer to the detector, counting the hits individually was no longer possible. A simple technique was developed to identify them

automatically. Firstly, the background was removed by subtracting the frame average from the field. Then the local maxima which had a value of at least 25 above the mean of the surrounding 3-by-3 pixel window were identified and judged to mark the point of maximum energy deposition for a particular electron hit. This process was verified by eye and found to be very good for both source energies and the resulting hit densities in the images.

Deconvolution procedures

The deconvolution procedures were similar to those used in the simulation section 4.3.1, with some differences to account for the changes in geometry.

The mask shape was the full repeated PBA mask and since the detector was a slightly different shape to this, it was necessary to select a 512×512 pixel region of the CCD frame. The mask shape was created using a Matlab script and resized to be the same scale as the detector - a total of 1024×1024 pixels. Cross-correlation of the cropped detector and the mask arrays would deconvolve the CCD output.

The mask image used for the deconvolution must be scaled when the source to be recovered is closer than infinity because of the diverging particle flux. For example, for a source at 16 cm from the mask, and a mask-detector separation distance of 4 cm, the mask shape must be expanded by 25% in each direction. For the radioactive sources, this was used to try to recover the source's spatial distribution. For the MCP-based source, the aim was to recover the angular distribution rather than the source position distribution, so for this the mask would remain at the same scale as the detector.

5.5.3 Results

A summary of the data produced from the CCD and the situations under which the data were taken is given here.

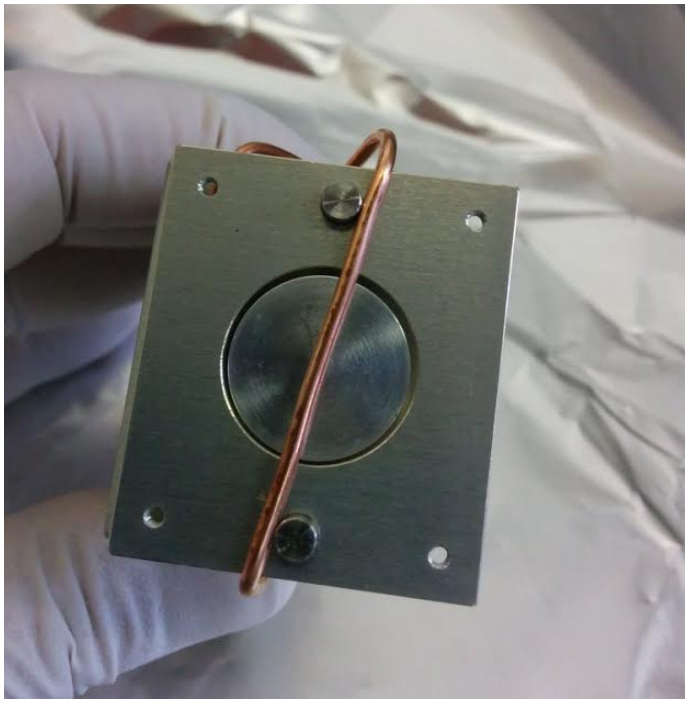


Figure 5.16: The samarium source with a copper wire across it

Radioactive source results

The initial frames were taken with a 2 mm diameter pinhole placed over the samarium source. This produced a very low flux indeed. Although the calculations given above would predict a particle count rate of $295\,\mathrm{s}^{-1}$, there were only 0 or 1 definite hits per frame measured.

A higher flux was attempted by opening the chamber, removing the pinhole mask from the front of the source, and creating a source shape with a greater total flux by leaving the surface open but placing a copper wire over the front of it. This would create a more complicated shape to reconstruct, but would give a chance at measuring a higher signal. This source configuration is shown in figure 5.16.

Although this setup produced a lot more detector hits, as expected, there were

still not enough for a deconvolution. The combined set of hits accumulated in a day's testing with a total exposure time of 4 hours is shown in figure 5.17.

Since the count rate was so low, the measurable count rate from this source was measured more accurately by moving the source to a position between the mask and the detector. It was mounted and aligned by eye approximately 1 cm from the detector surface. This would be close enough that the activity per surface area of the source would be expected to match the number of hits on the CCD.

One of the 100 s exposures with the source placed approximately 1 cm in front of the detector is shown in figure 5.18, scaled so that as many hits as possible are visible. However, there were obviously far fewer hits than would be expected for the quoted activity.

Since the count rate was insufficient using the samarium source, a carbon-14 source was used in its place. Unfortunately, this source had a similarly low count rate for tests both with a full disc exposed and with a 2 mm diameter pinhole.

As a demonstration of the flux of the carbon source and the ability to detect it, exposures were made of carbon-14 from a distance of approximately 1 cm from the detector. The result of this is shown in figure 5.19. Again, the count rate was far lower than expected. Since angular measurements were not possible with the radioactive sources, the decision was made to use the adapted MCP-based source.

MCP source results

The images taken with the MCP source produced far more complex data than the radioactive source setups. Figure 5.20 shows a series of representative frames. These images show an HV supply set at 2000 V above the total voltage selected by the Zener diodes, so the voltage across the MCP was expected to be 2000 V. This aimed to produce an equal number of electrons at each of 0 eV, 400 eV, 800 eV, 1200 eV, 1600 eV and 2000 eV. The intensity measured on the detector will be compared in section 5.5.5 with the energy deposition for a single electron predicted by the simulations.

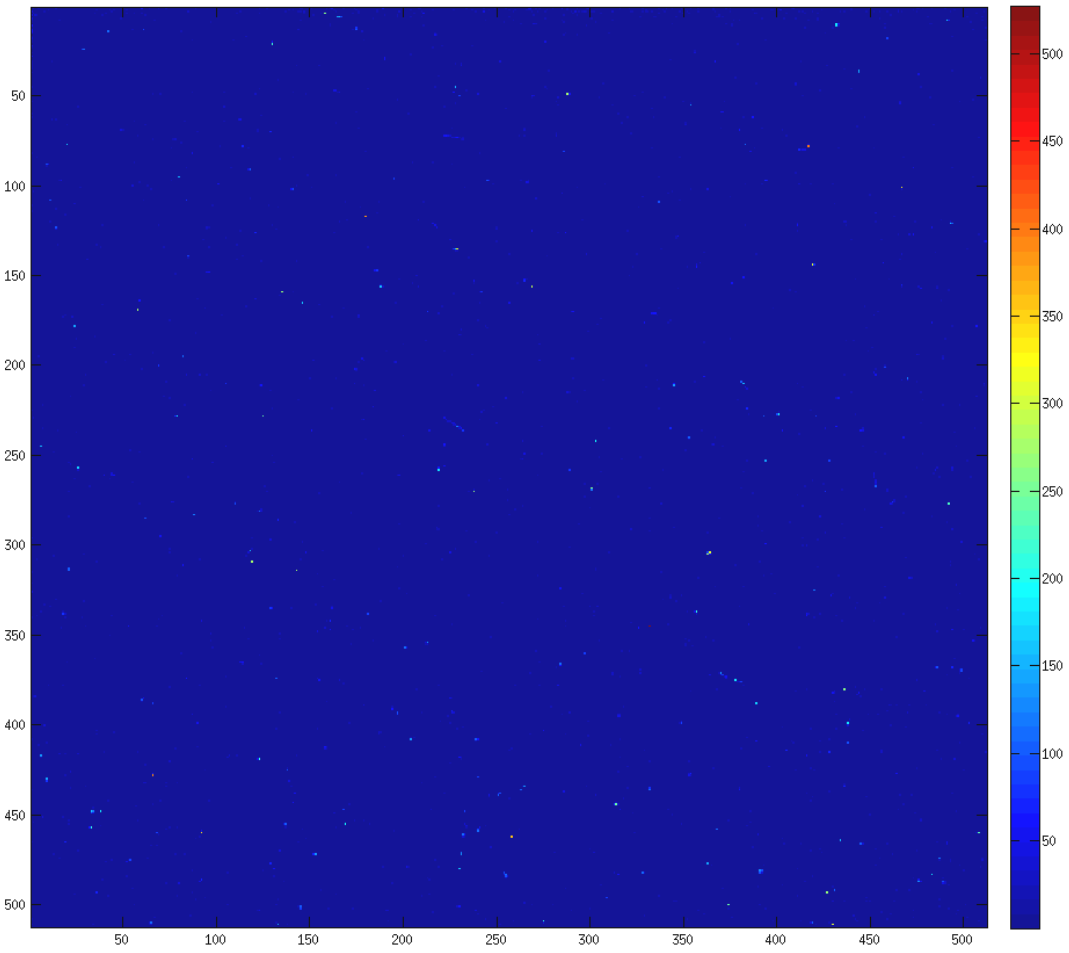


Figure 5.17: Summed hits (bright points at least 20 DN above local background) for 143 out of the 155 frames of 100 s exposure taken on 18 November 2014. The source was the full disc of samarium-151 with a copper wire across it, at approximately 16 cm from the detector.

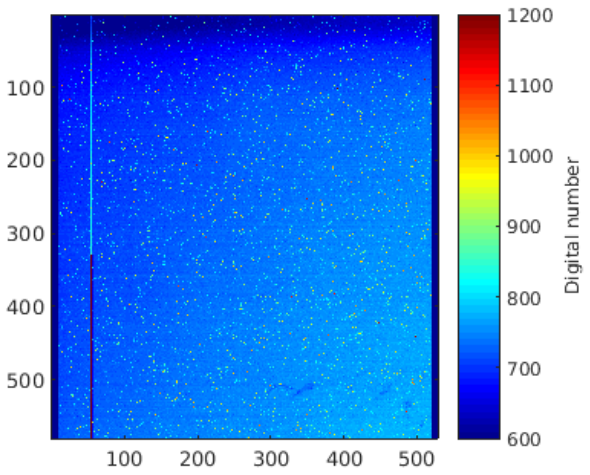


Figure 5.18: The digital numbers from each pixel from a single 100 s exposure with the samarium source approximately 1 cm from the detector.

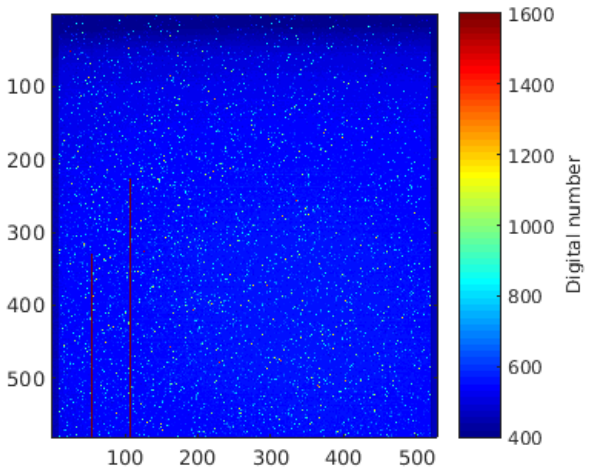
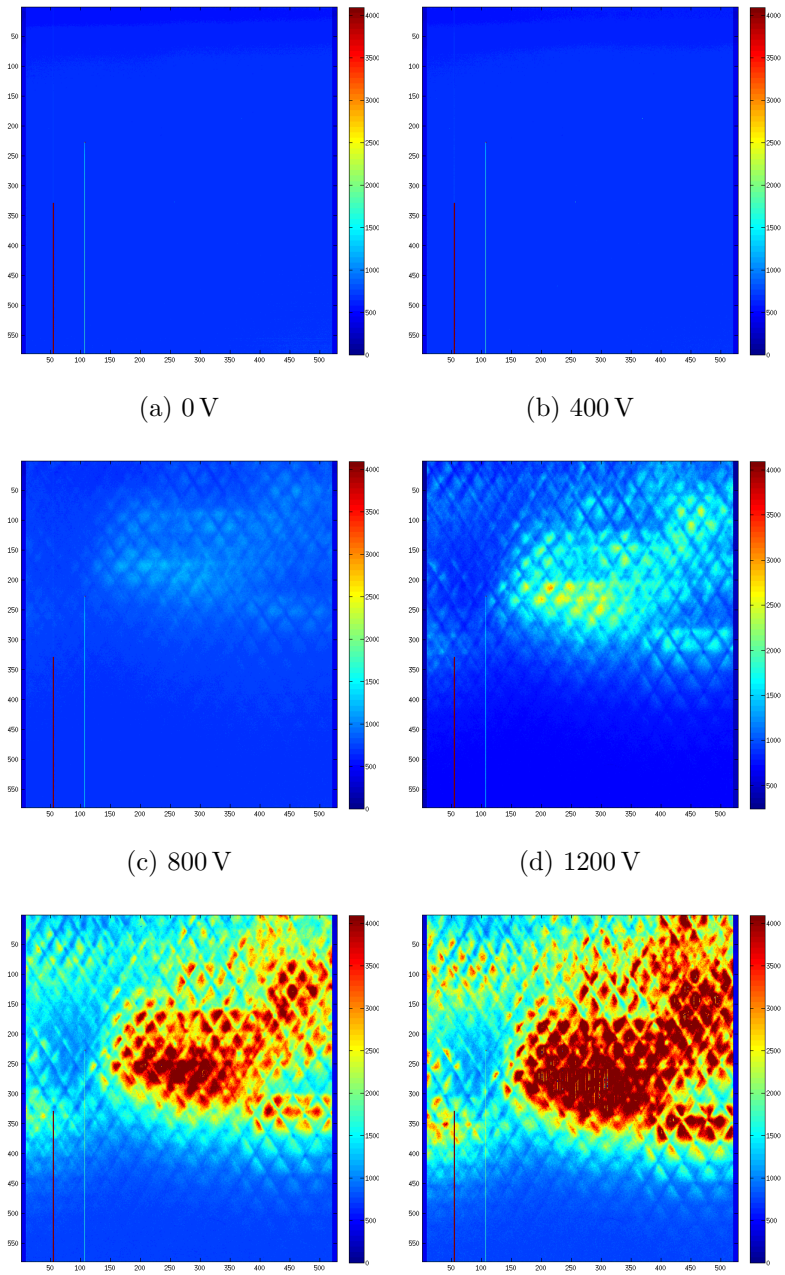


Figure 5.19: A 100 s exposure frame with the carbon-14 source positioned approximately 1 cm from the detector.



(a) 0 V (b) 400 V

(c) 800 V (d) 1200 V

(e) 1600 V (f) 2000 V

Figure 5.20: CCD frames with the mask in place and 2 kV across the MCP at each acceleration voltage.

In addition, a series of frames were taken with the mask removed. These provided a background distribution to highlight any position sensitivity of the CCD and variation of the MCP flux in the detector plane. An image frame is shown in figure 5.21 for 2000 V acceleration voltage. Since without the mask in place, the number of particles reaching the detector was lowered, the voltage across the MCP was reduced to 1800 V to avoid the CCD saturating in the 100 s exposure. The diagonal pattern in the image is a result of the CCD manufacturing process and will be discussed in section 5.6.4. No frames had been taken with the mask in place at 1800 V across the MCP since the pattern on the CCD frame became very faint at 1900 V.

Although these results show a great deal of structure, attempting to deconvolve these images was unsuccessful. Different scale factors and offsets of the mask were tried to visually identify the mapped position of the mask shape. One of the attempted mask overlays is shown in figure 5.22, although this is clearly not sufficient to explain the measured pattern, since large areas of the detector are not matched.

Section 5.6.4 will cover the efforts to understand the MCP data using a mixture of analysis and simulations.

5.5.4 Noise sources and levels

Expected causes of background and noise included thermal noise, readout noise and other electronic noise. Part of this noise was related to exposure temperature, and part was, in good quality frames, reasonably constant for a given exposure time.

The average of mean frame value against temperature is shown in figure 5.23 for frames which were both empty of signal sources and of clearly good quality. A best fit line was plotted according to the dark current temperature dependence stated in the CCD64 datasheet, which is proportional to $T^3 e^{-6400/T}$, and an offset which might be attributable to effects produced by the readout and amplification electronics.

For background scaling and removal, the coefficients of this fit line were used for

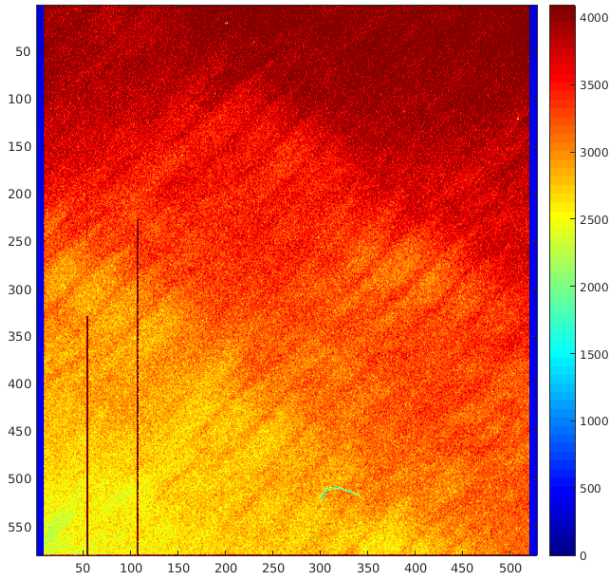


Figure 5.21: A CCD frame taken without a mask and with the MCP source still in place. The exposure time was 100 s. In this case, 1800 V across the MCP and 2000 V between the front of the MCP and the grounded grid. Reducing the MCP voltage from 2000 V to 1800 V was necessary to stop the CCD saturating. The diagonal structure on this image is due to the structure of the detector and is unrelated to the mask pattern. It will be explained in section 5.6.4.

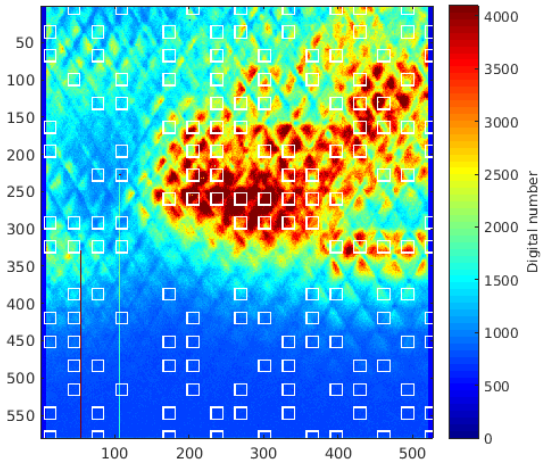


Figure 5.22: An attempt at matching the mask pattern to the detector output. The white outlines show a possible position of the mask over the CCD.

the few warmer frames where the temperature change had a noticeable effect. The best fit curve gave a constant expected offset of around 696 ADU without dark current. This fit also demonstrates that frames taken where the detector was measured below approximately −20 °C had relatively little dark current. Since most frames were taken below this temperature, the majority of the background was attributed to effects in the readout and amplification electronics.

5.5.5 Detection of electrons using a CCD64

A CCD64 has previously been used for the detection of low-energy electrons, as described by Bedington et al. [2012]. To improve upon the current simulations of low energy electron behaviour in this sensor, the CCD was simulated in Geant4 as a thin dead layer of silicon dioxide and an active region of 15 µm of silicon. Based on these previous calculations and measurements, the dead layer was assumed to be approximately 5 nm thick.

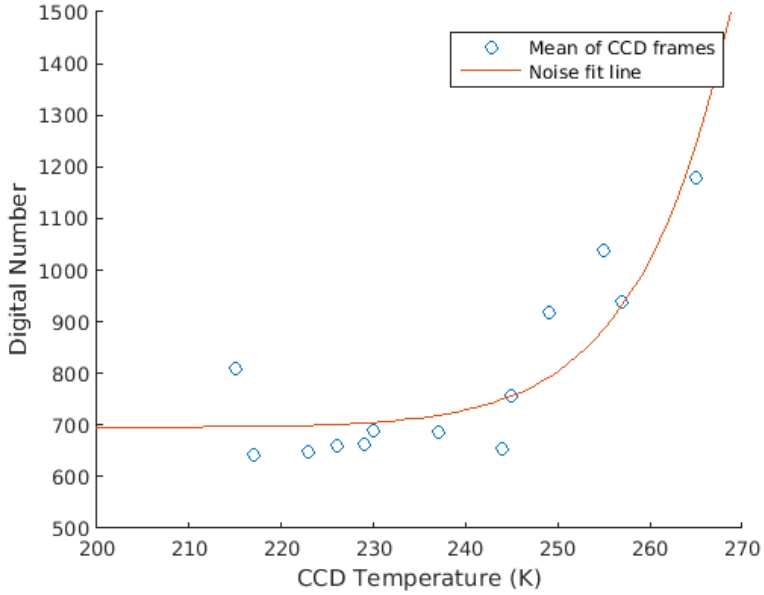


Figure 5.23: A selection of the average pixel value of 100 s exposures to assess the dependence of background on temperature. About half of the frames were judged by eye to have sufficient quality to be included.

Energy deposition and depth

For each electron energy between 100 eV and 2 keV in 100 eV steps, ten thousand electrons were flown in Geant4 from just in front of the simulated detector surface, through 5 nm of silicon dioxide and a thick layer of silicon. Since this simulation required analysis at very low energy and distance scales, the cut range and step function resolution were both reduced and the Livermore ionisation model was used, which is a low-energy electron model. These calculations are similar to the calculations by Bedington et al. [2012] who used Mulassis simulations with varying layer thicknesses, but give an improvement in resolution.

The energy deposited in the detector was recorded in a CSV file as in earlier simulations, but binned in 0.2 nm depth bins and the resulting energy deposition against depth is shown in figure 5.24.

In order to translate these values into the amount of energy absorbed and measured by the detector, an estimate of the CCE and its dependence on depth in the detector must be made. Semi-empirical models have been made, including those by Stearns and Wiedwald [1989] and Stern et al. [1994]. In these models, the CCE decreases away from the collection electrodes towards the surface of the detector.

The CCE described by Stearns and Wiedwald [1989] was valid beyond the dead layer, where the efficiency was zero:

$$\eta\left(x\right) = 1 - \left(1 - \eta_0\right) e^{-x/d} \tag{5.1}$$

where η_0 and d must be determined by experiment.

Using the coefficients from Stern et al. [1994] who used an ion-implanted back-illuminated CCD comparable to the CCD64 and an assumed 5 nm dead layer, this CCE is given by

$$\eta\left(x\right) = \begin{cases} 1, & x < 5\,\mathrm{nm} \\ 1 - 0.79 \cdot e^{-x/230\,\mathrm{nm}}, & x > 5\,\mathrm{nm} \end{cases} \tag{5.2}$$

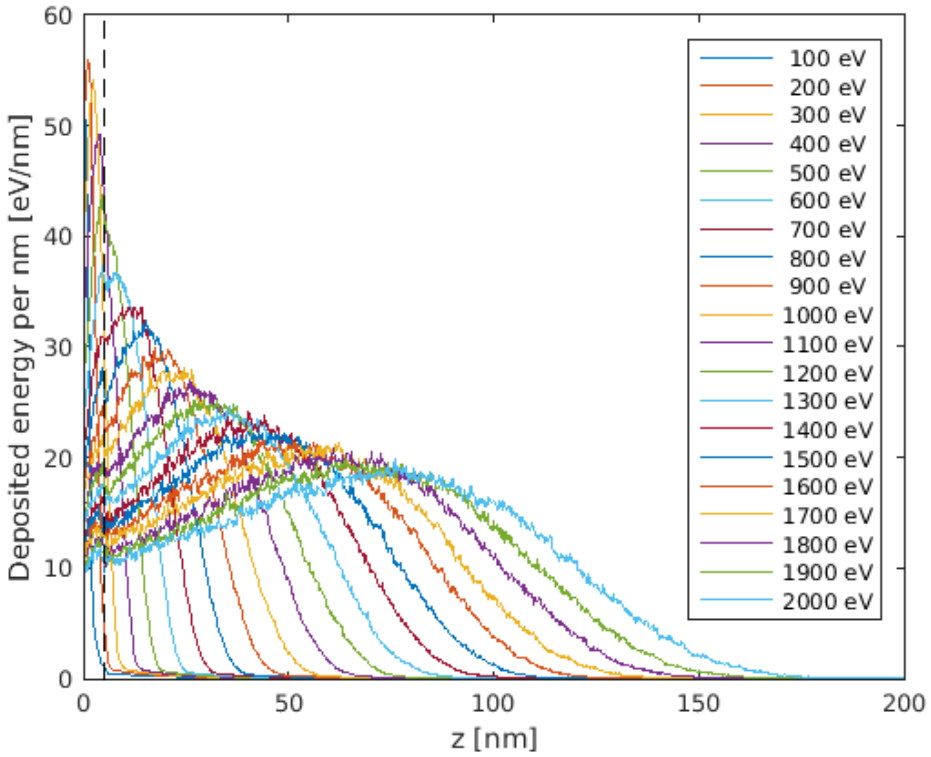


Figure 5.24: The energy deposited by electrons from 0 eV to 2000 eV energy at each depth. The 5 nm region marked at the left is the assumed silicon dioxide dead layer.

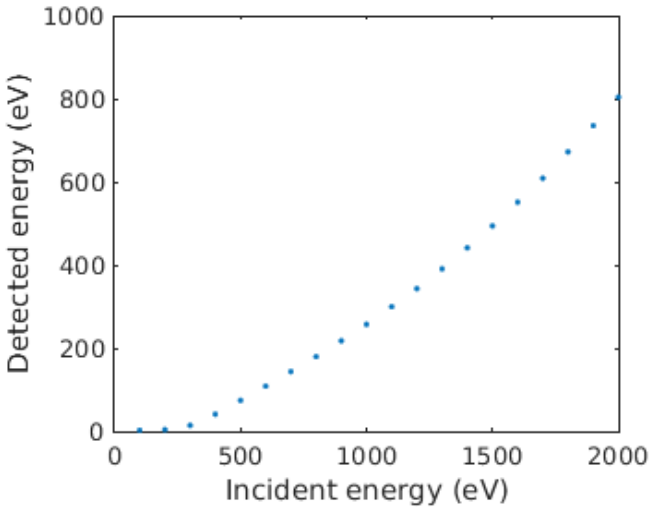


Figure 5.25: The detected energy of an electron with incident energy up to 2 keV based on Geant4 simulations

where x is the distance from the surface of the detector.

Combining the data in figure 5.24 with equation 5.2 produces the average collected energy from an electron at each energy. This is shown in figure 5.25. This provides a scaling factor for the intensities in the frames taken with the MCP-based source.

MCP source flux

These calculations were used to make a first estimate of the number flux of electrons reaching the detector from the MCP-based source. The average pixel values in the active region for each undamaged frame taken in one day's run of experiments is shown in figure 5.26, and can be seen to be highly repeatable for each MCP voltage.

The signal in the detector for each energy can be compared using the expected detected energy shown in figure 5.25. Using the conversion of 3.65 eV of deposited energy to create a generated electron-hole pair, and the stated gain of the electronics,

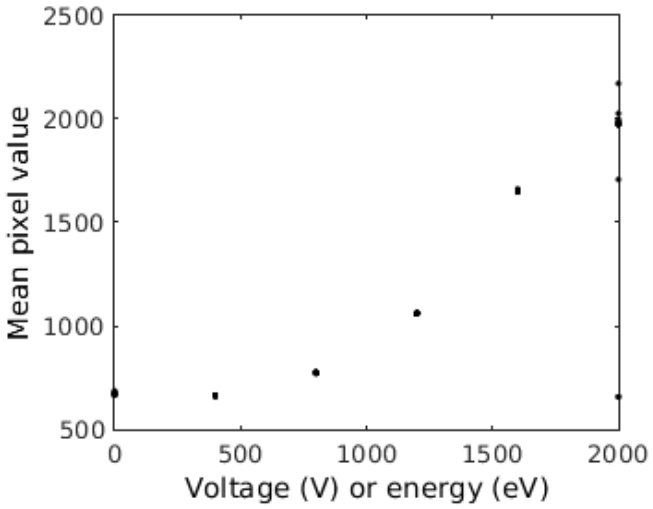


Figure 5.26: The mean pixel value for all undamaged 100 s exposures taken on 20 February 2015

where 25 electron-hole pairs produce one ADU, an estimate of the detected number flux of each energy of electrons can be given. No signal is measured with MCP acceleration voltages of 0 V and 400 V, but the average of frames between 800 V and 2000 V give estimates of 56, 106, 164 and 149 particle hits per pixel.

There is a significant number of saturated pixels in the output from the 2000 V runs. It is also apparent that the outline of the shape seen on the detector is shifted slightly depending on energy, indicating that the angular distribution varies with the particle energy. Although the MCP source was manufactured and designed to have a zero bias angle, its exact distribution was unknown, as well as the dependence of the final angle on the voltage across the acceleration plate. These may explain part of the reason for the variation in the apparent number of particle hits.

Point spread functions

The properties of the PSF of an electron interaction are important, both for identification and positioning of a hit and for an estimate of electron flux based on accumulated charge, depending on the mode in which the CCD is operated.

The PSF of an electron hit will depend on the motion of the electron within the silicon and on the charge spreading once the electron hole pairs are produced. In the case of lower energy electrons, including the majority of electrons used in these lab-based tests, the electron will quickly run out of energy and most of its energy will be deposited near its entry point.

Simulating the CCD with an assumed 5 nm dead layer of silicon dioxide and 15 µm of silicon active region can give the expected probability density for energy deposition in each pixel surrounding the electron entry point. The simulation of ten thousand electrons at each energy gave the expected energy deposition in the active region per incident particle per pixel shown in figure 5.27 with a log scale to show both low and high pixel values.

This can be compared to the sample areas of the data from the CCD shown in 5.28. These have the background subtracted and the particle hits are clearly visible. The Q-value of carbon-14 is 156 keV and the Q-value of samarium-151 is 77 keV. It can be seen that many of the particle hits in the carbon-14 frame form tracks across several pixels. The samarium-151 electrons hits usually produce a signal in one or two pixels, as expected.

In a running mode where the CCD is left to accumulate hits over a long exposure time, individual hits and PSFs are not identified, but the distribution of the response to hits of a certain energy should be known in order to understand the resultant smudging of the mask's shadow pattern. For the MCP source, this accumulation was measured, but at these low energies, the electron was almost always expected to produce a signal in one pixel only and the PSF was not a source of smearing in the image.

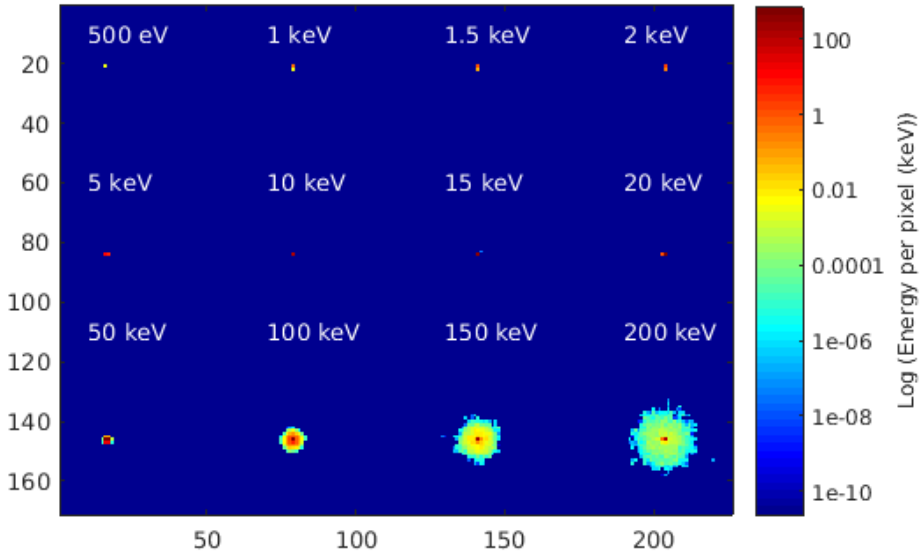
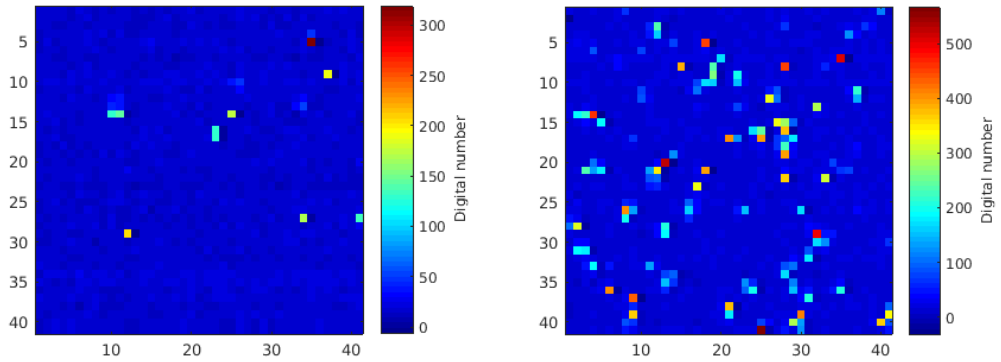


Figure 5.27: The simulated probability density of energy deposition of an electron of each energy in the CCD64 with a logarithmic scale.



(a) Samarium-151 (b) Carbon-14

Figure 5.28: A 40×40 pixel area of the detector frame for 100 s exposures with tests with each source close to the detector.

Energy spectra

An estimate of the impinging particle energy spectrum was produced for each radioactive source used, based on the detector images produced of the electron hits when the source was placed close to the detector.

Using the hit identification method described in section 5.5.2, a single frame with 100 s exposure time from the samarium-151 test was found to contain 7739 particle hits, and the equivalent from a carbon-14 test contained 10646 particle hits.

The energy deposited by a single particle was estimated by summing the background-removed 3×3 window surrounding the maximum point identified for each particle hit. The distribution of these summed ADUs is shown in 5.29. The distribution spectrum of samarium electron energies is closely related to the expected beta-distribution, while the carbon emission is slightly less as predicted. This is likely to be for several reasons. Firstly, the higher-energy electrons in the carbon-14 spectrum were more likely to produce a signal outside the 3×3 window that was summed and secondly, these higher-energy electrons were likely to reach the back of the active

part of the detector and only part of their initial energy was measured by the CCD. In addition, the higher count rate and wider PSF were likely to lead to more overlapping hits. This effect can be seen in a small second peak at around 420 ADU. In both source cases, the noise in the detector contributed to the uncertainty in the deposited energy. The expected beta spectra and simulations of them are described in detail in section 5.6.1.

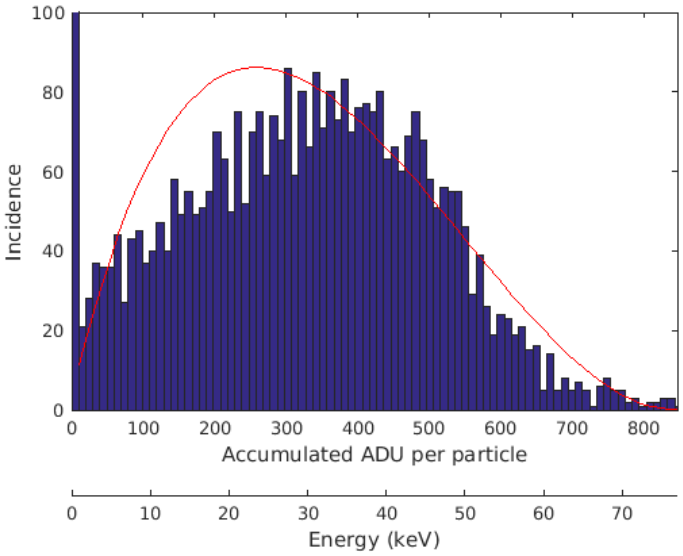
The energy measured in accumulated ADU is related to the incident electron energy. The Q-value of samarium-151, where the emission spectrum ends, is 77 keV. If an electron with an energy of 77 keV deposited all of its energy in the active region of the detector, it would produce 21095 electron-hole pairs, which with the stated gain of 25 would produce an accumulated digital signal of 843 ADU. This matches the estimated beta spectrum in the data well. The carbon-14 spectrum is more complicated, since the PSF is wider and a small deposited energy could be caused by either a low energy incident particle or a high energy one that moved rapidly through and out of the active region.

The differences between the energy spectrum measured by the CCD and the calculated beta spectrum equation will be revisited in detail in section 5.6.1. However, these measurements give a good indication that an instrument based on a CCD64 is capable of energy measurements when the flux and measurement cadence are matched to give hit numbers of this order of magnitude and the incident particles are known to be within a certain range.

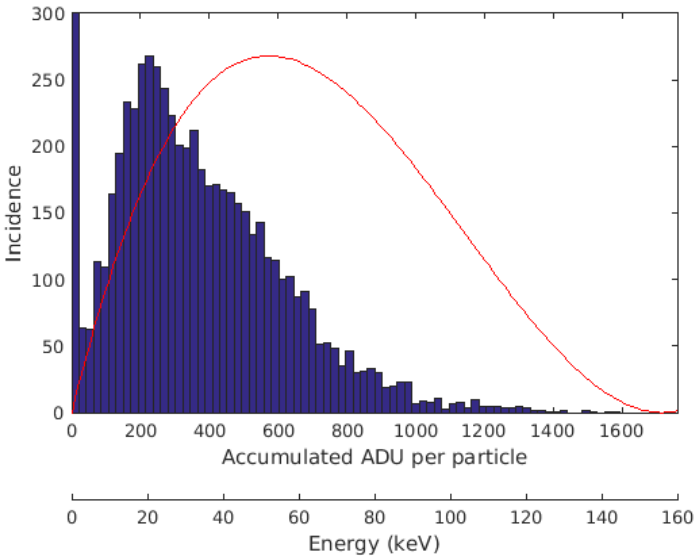
5.6 Comparison with simulations

The lab setup was replicated to a basic degree in the Geant4 environment to further understand the collected data. This required the building of the detector, mask and source in a Geant4 application. Other parts of the experiment setup had relatively little impact on the output of the experiments and were not simulated.

In each of these simulations, the CCD was simulated as an active region of



(a) Samarium-151



(b) Carbon-14

Figure 5.29: The estimated energy spectra in deposited ADU measured from the tests with each source placed close to the detector. The red line shows the predicted beta spectrum calculated using the method described in section 5.6.1.

15 µm of silicon, with a very thin dead layer of 5 nm of silicon dioxide, based on the estimates used by Bedington et al. [2012]. The total area of the detector was 8.4 mm×9.3 mm, which in the Matlab processing was pixellated to 528×580 16 µm×16 µm pixels.

The mask was modelled as 470 µm thick copper, with a repeated PBA geometry with a total aperture open fraction of 12.5% and hole size of 256 µm as used in the experiments.

5.6.1 Approximation of the energy spectra of radioactive beta sources

For the first set of lab-based instrument tests, the source of energetic electrons was β-decay - where a nuclear neutron decays into a proton, a neutrino and an electron. Since there are three particles created, and only two constraints on their energies and velocities (conservation of energy and momentum), the electron velocities are isotropic and their energies are distributed within a beta emission distribution.

For many radioisotopes, the spectrum of their decay products has been measured by experiment, but several analytical forms have been developed for the general case. Equations for a fit to this energy spectrum are useful as an input to simulations.

An early equation for the probability of the emission of a beta spectrum was published in a paper by Fermi [1934] which was translated from the Italian by Wilson [1968]. This probability distribution function is perturbed by the Coulomb field of the produced daughter nucleus, so a correction factor is added to account for this effect. This correction also takes into account forbidden transitions and the screening effect of orbital electrons. The total emission energy distribution function for an allowed transition can be calculated using

$$N(E) = CF(Z, E)p^2(Q - E)^2, \tag{5.3}$$

from, for example, Cohen [1971], where C is a constant at relatively low energies,

F is the Fermi distribution function equivalent to the total correction factor, Z is the atomic number of the nucleus, E is the kinetic energy of the particle, p is the momentum of the particle, and Q is the upper bound of the energy spectrum (determined by experiment).

The Fermi function is complicated to define analytically, Feister [1950] described some of the approximations that have been developed, with adjustments for forbidden transitions. For example, a useful approximation for the Fermi distribution function for low Z-values up to 29 and low kinetic energies was described by Kurie et al. [1936]:

$$F(Z,\eta) = \frac{2\pi y}{1 - e^{-2\pi y}}. \tag{5.4}$$

where

$$y = \gamma \frac{(1-\gamma^2)^{\frac{1}{2}}}{\eta},$$

$\gamma = Z\alpha$ where α is the fine structure constant, and

η is the momentum of the electron.

For higher values of Z, there have been multiple more advanced approximations, for example those by Hall [1950]. The expression used by Bethe and Bacher [1936] was given in a simplified form by Longmire and Brown [1949] which is judged to be accurate to about 1% up to $Z = 84$. This is given by

$$F\left(Z,\eta\right) \approx \frac{2\pi y}{1 - e^{-2\pi y}} \eta^{2S} \left(y^2 + \frac{1}{4}\right)^S \tag{5.5}$$

where $S = (1 - \alpha^2 Z^2)^{\frac{1}{2}} - 1$.

The Q value of samarium-151 is 77 keV and the atomic number is 62. Therefore to calculate its spectrum the more complex approximation in equation 5.5 was used. The distribution calculated according to the above equations is given in figure 5.30 (a). This shows the most frequent electron energy is approximately 23 keV. The polynomial is a fit to the energy distribution, which was used in the simulations

of the setup.

The Q value of carbon-14 is 156 keV and the atomic number is 6. In this case the Fermi function approximation given in equation 5.4 was adequate. This curve, and a polynomial fit to it, are shown in figure 5.30 (b).

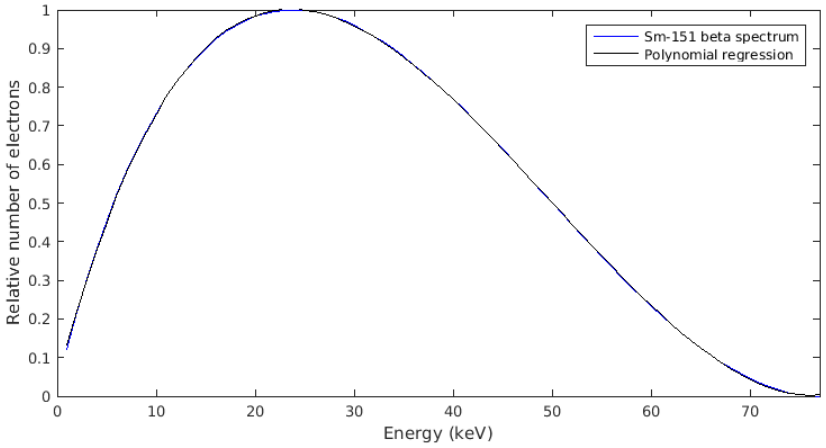
These spectra were used to define the energy distribution of the electrons used in the simulations.

Comparison of simulated energy deposition against measured

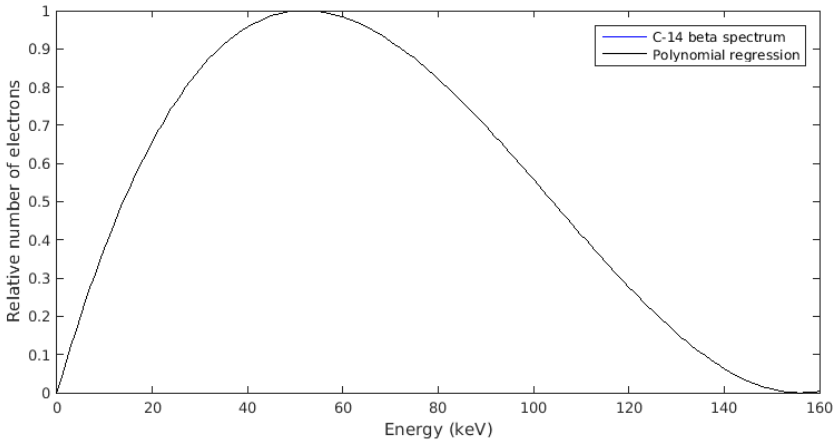
To investigate the difference between the calculated expected emission spectra and the measured spectra in the experiments, a set of simulations of these two radioactive sources was run. This was to understand the influence of the geometry of the detector and the algorithm used to identify hits in the CCD on the shape of the measured spectrum, and to explain the differences between the shapes of the calculated and measured spectra.

For these simulations, a model of the CCD64 geometry was used. A large number of electrons with an energy spectrum from the polynomial fitted to the above calculated curve was flown. The deposited energy was then binned in CCD-size pixels and multiplied by approximately 11 ADU/keV to produce a similar frame to the same scale as the CCD64 frames from the experimental setup. A cutoff was chosen for the simulation data so that the number of particle hits was the same for both the simulated and experimental data. For one frame of each of simulated and experimental, the algorithm from section 5.5.5 was used to produce a measured energy spectrum.

Figure 5.31 shows the three normalised distributions together: the calculated beta emission spectrum of each source as in figure 5.30, the measured energy spectrum from experiments as in figure 5.29 and the same spectrum recovery algorithm on the simulated CCD frame. The measured and simulated curves match well, explaining that almost all of the disparity between the predicted spectra and the spectra recovered from the CCD frames is a result of the geometry of the detector.

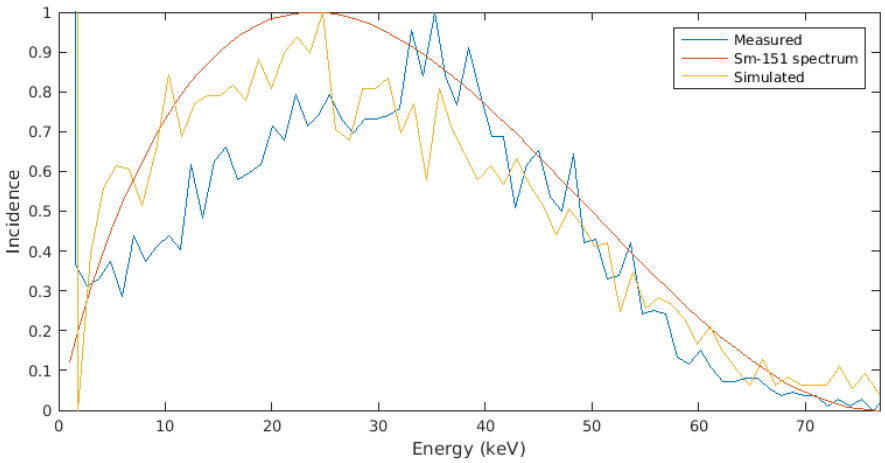


(a) Beta-emission spectrum of samarium-151

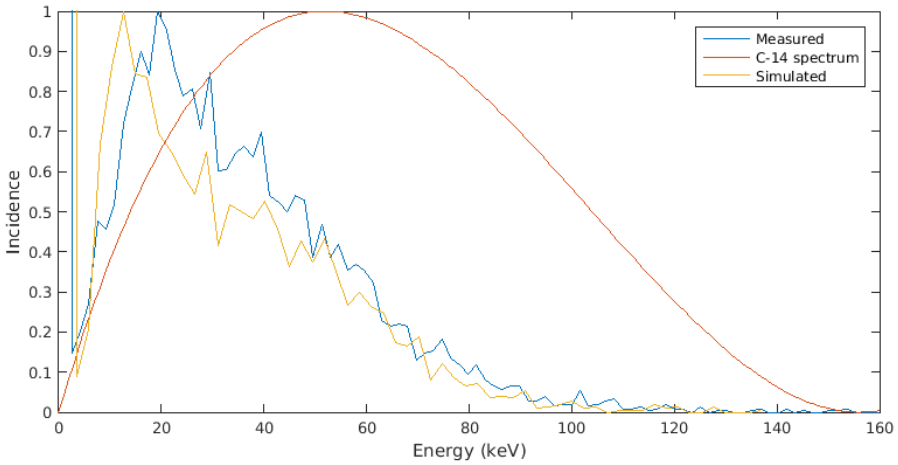


(b) Beta-emission spectrum of carbon-14

Figure 5.30: Beta-emission spectra based on the calculations in equation 5.3. The blue line indicates the calculated values. The black line is the cubic polynomial fit used in the simulations. In both cases, the fit is very close to the calculated spectrum.



(a) Samarium-151



(b) Carbon-14

Figure 5.31: Comparison between measured (as shown in figure 5.29), calculated (as shown in figure 5.30) and a simulation with a model of the CCD64 with an incident particle spectrum based on the previous calculations.

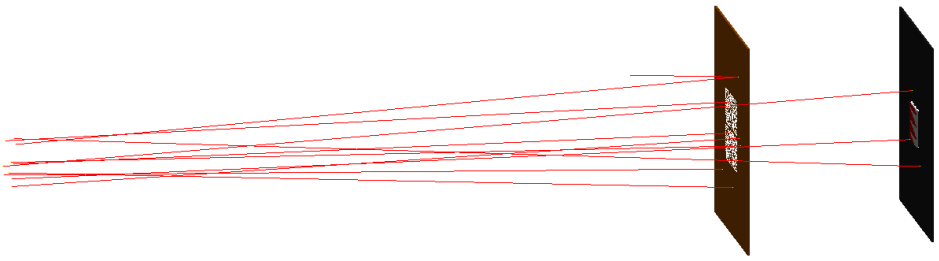


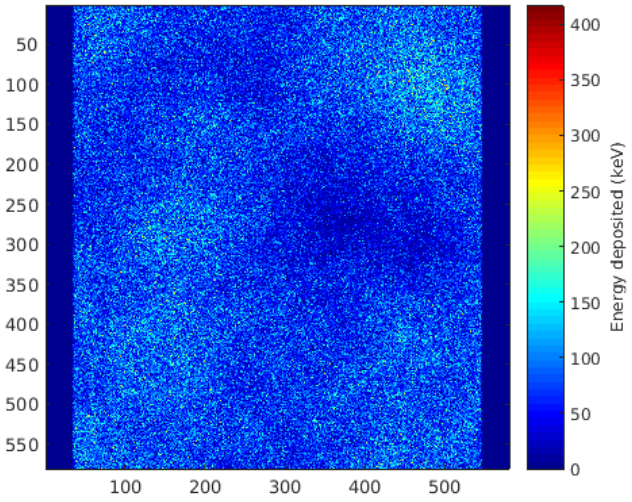
Figure 5.32: An image of the setup of a carbon-14 source simulation showing the source, detector, and 10 beta particles within a narrow range of angles

5.6.2 Simulation of the radioactive source experiments

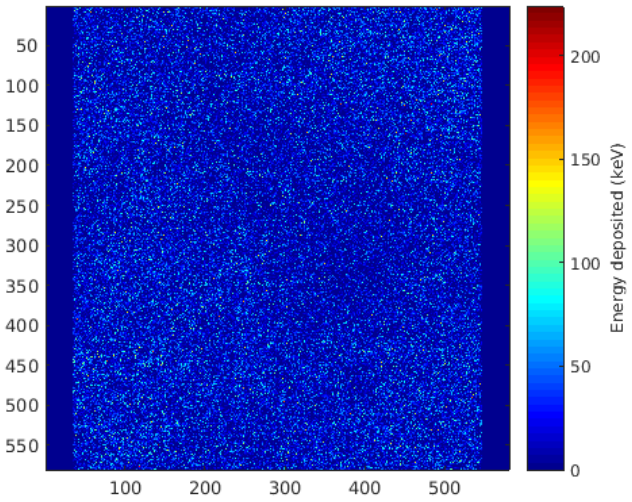
The position and angular distribution of the simulated beta particles was based on the size of the source used, although only those particles aimed in the direction of the detector were simulated to improve the speed and efficiency. The number of particles flown was based on 100 s of the stated activity level emitted isotropically. An example of the setup in Geant4 is shown in figure 5.32.

The output of the simulation of a 100 s exposure of each source is shown in figure 5.33. A limited amount of direct comparison was possible. Visually, the output frames look far more similar to the experiments run with the source 1 cm from the detector than those run with the mask in place.

Using the routines to identify and measure particle events in the experiment data, the number of particle hits on the detector in this simulated experiment was calculated. For the samarium source, the number of counts in a full frame of a simulated 100 s exposure was 31475 counts, and for the carbon source it was 24335 counts. This is significantly higher than the single-digit number of hits in the real experiment. The disparity between these values based on the stated activity and



(a) Samarium-151



(b) Carbon-14

Figure 5.33: Simulated 100 s exposures

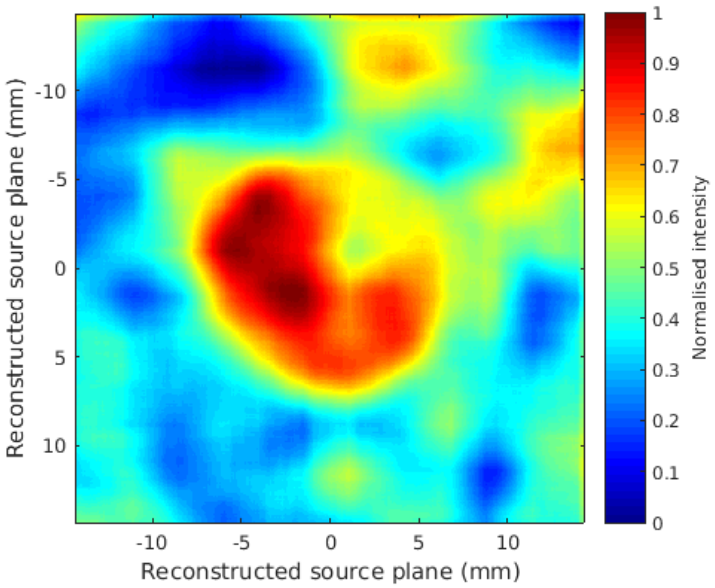


Figure 5.34: The deconvolved data from figure 5.33 (a). The image is normalised to between 0 and 1.

the experiment values might be due to the way the activity was measured, damage to the source or the way the source mounting was made.

With a far larger number of hits than in the experimental data, a deconvolution was possible and is shown in figure 5.34. The mask image was scaled up to account for the finite difference of the source. Although the image quality is not particularly good and the SNR is only approximately two, a circular source can be reconstructed, and a diameter of 14 mm to 15 mm can be seen. Further exposures or a longer exposure time would improve this image. Although CCD data from experiment would have some noise component in addition to this signal, the technique is relatively resilient to noise, as demonstrated in section 4.3.7. For another disc-shaped samarium source with higher activity, the ability to reconstruct such a distribution would be expected to be possible.

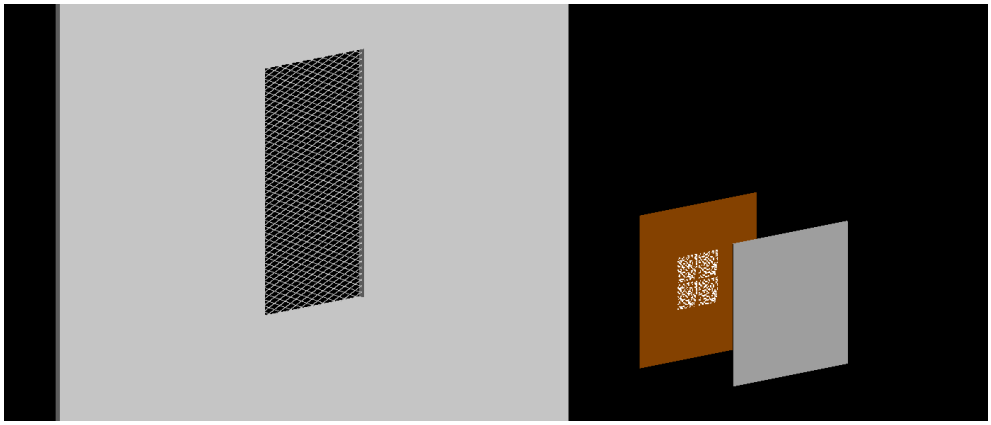


Figure 5.35: The model of the MCP source in Geant4 showing the grid shape of the grounded electrode at the left.

5.6.3 Building a model of the MCP-based source

To begin with, much of the behaviour of the MCP source was unknown. Any previous measurements of the source behaviour had been made using an uncalibrated instrument prior to the modification of the switchbox. The grounded grid at the front surface of the MCP source could be seen to have a diamond pattern, but beyond this, the precise configuration of the distribution of particle positions or angles was uncertain.

An attempt to produce a similarly-patterned electron source based on what could be discovered during these tests was made, and a model for the source was produced. The MCP was built in Geant4 as shown in figure 5.35. This was a plate of aluminium with a 75 mm by 39 mm window with a grid which was also made of aluminium. Each of the wires in the grid was 200 µm thick. In this case, the angles of the grid wires were at 30° from the horizontal in each direction.

The simulations were set up as shown in figure 5.36. The source particles were created directly behind the grid. The angular distribution of the electrons produced was to be determined iteratively by simulation.

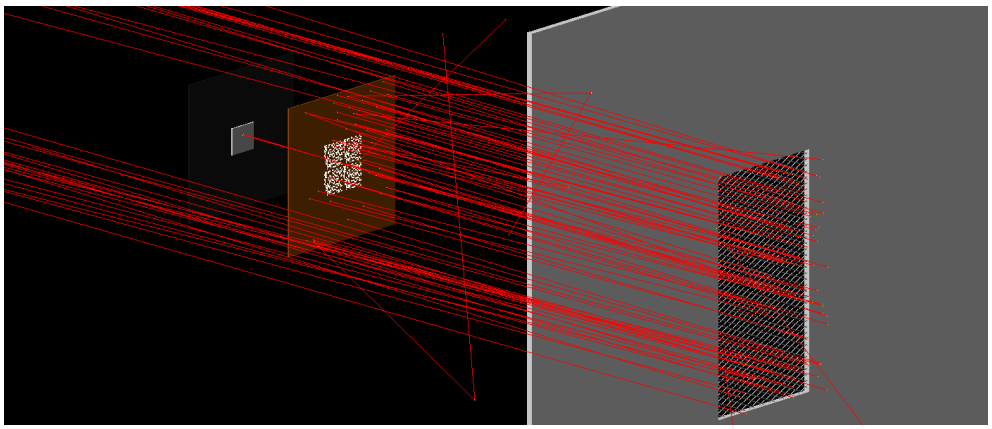


Figure 5.36: The setup of the MCP simulations, with an evenly-distributed source of 2 keV electrons situated behind the MCP grid.

5.6.4 Simulations to attempt to understand the MCP source

The results from the tests with the MCP source were, at first, confusing. The detailed structure was unexpected. The diamond pattern evident in all frames from the MCP with the mask is present, and shown again in figure 5.37.

This diamond structure was initially considered as possibly derived from structure on the CCD itself, as a result of the grinding process used in back-thinning the wafer from which it was made. These patterns had been measured for a very similar CCD and described in Stern et al. [2004]. The data taken from illumination with a blue LED showing the grinding marks were provided by D. Walton. This pattern was verified for the MCP electrons when background measurements were taken without the mask. This full illumination of 800 eV electrons is shown in figure 5.38. This effect did not well match the diamond pattern and was not sufficient to explain the MCP source readings.

The next most obvious theory for the diamond pattern seen in the image frames was that it could be a projection of the diamond pattern of the grid on the MCP source, which might have been imaged through each pinhole. The shape of the

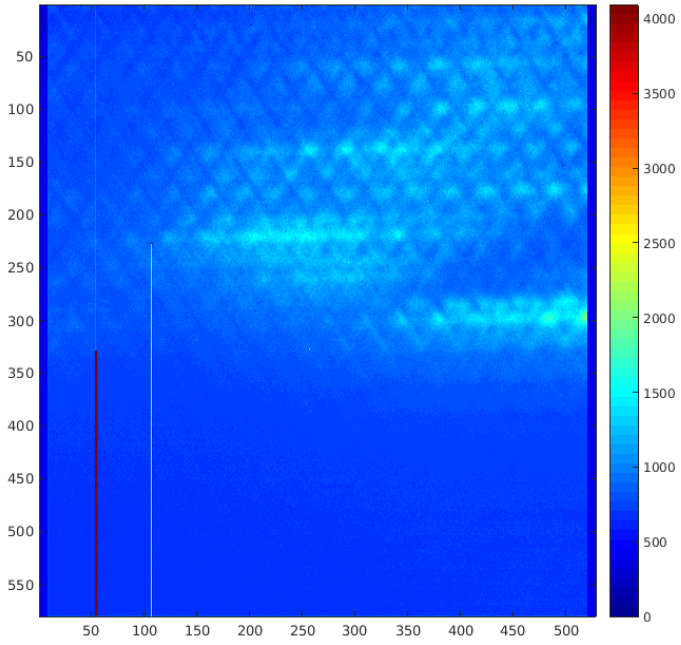


Figure 5.37: A 100 s exposure CCD frame with the mask in place, 2000 V across the MCP and 800 V acceleration voltage showing the diamond-shaped structure in most of the image.

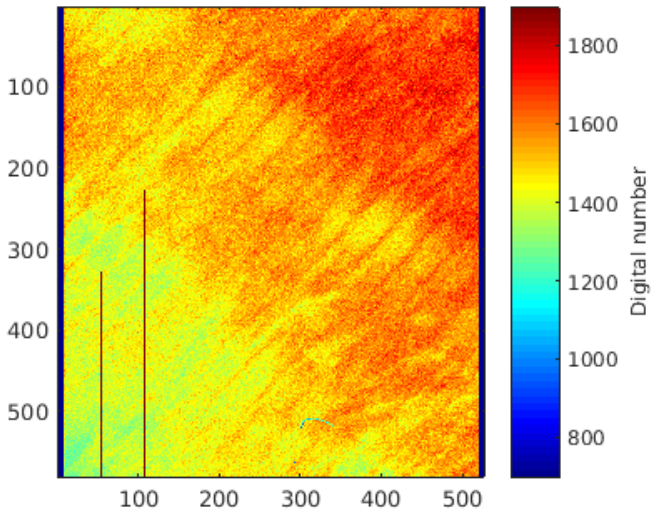


Figure 5.38: A measured near flat-field exposure without a mask of electron energies of 800 eV with 1800 V across the MCP and a 100 s exposure time

imaged diamonds corresponded to the shape of the grid in front of the MCP, and the angular alignment was correct also. However, the measured sizes of the imaged diamonds on the CCD did not match this hypothesis: they were approximately 67 pixels, or 1.07 mm, long and 40 pixels, or 0.64 mm, wide. For a designed mask-detector distance of 4 cm and a measured source-mask distance of 17 cm, this would be an image of a grid pattern of diamonds on the MCP source face of 4.55 mm by 2.72 mm. This is significantly larger than the diamonds pattern on the face of the MCP source in reality: 3.25 mm by 1.85 mm measured with a ruler on the MCP.

To further check this hypothesis, the mask in the simulation was replaced with a single pinhole; the image produced by each single hole in the mask would be examined. The initial simulation run was exceptionally slow (clearly demonstrating the improvement in sensitivity using a mask rather than a single pinhole), so the speed was improved by simulating only those particles that were aimed at or around the detector. Figure 5.39 shows that such simulations have no visible diamond-based pattern for this size and shape of pinhole. This indicated that the MCP-based source must produce a more complicated position and angular distribution of electrons.

After some trial and improvement, a series of angular distributions which produced simulated image frames which qualitatively matched the patterns in the data frames was found. An example of the angular distribution of one of these simulated CCD frames is shown in figure 5.40. This should be compared with the real data frames shown in figure 5.20.

The source distribution used to produce this frame using the mask was an even distribution of 2 keV electrons behind the simulated MCP grid with an angular distribution given as shown in figure 5.41 (a). Again, this simulation was sped up by choosing not to simulate particles that were not aimed in or close to the direction of the detector.

It was shown that this diamond-based pattern could be reconstructed from the output shown in figure 5.40 in figures 5.41 (b) and (c), using two different resolutions of the mask pattern.

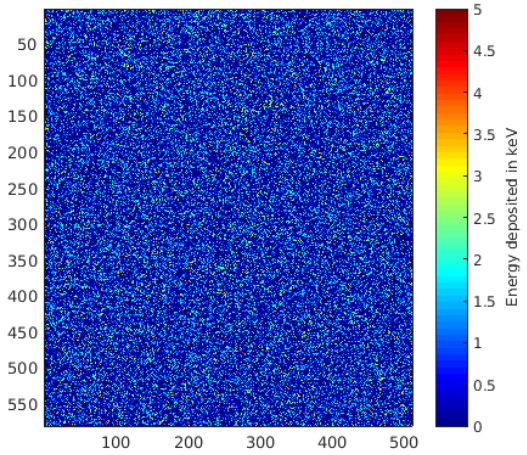


Figure 5.39: An omnidirectional 2 keV source behind a model of the MCP source grid, using a mask with a single hole to show the PSF of an expected source pattern.

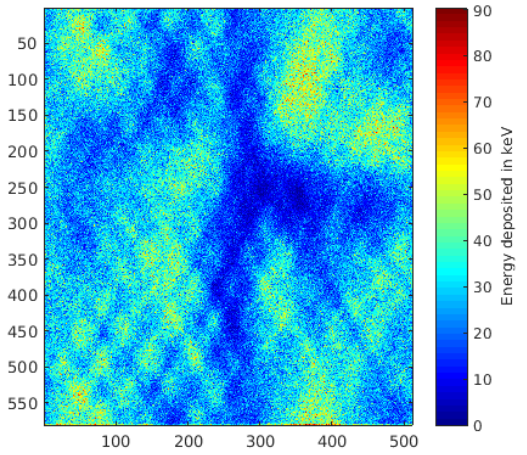
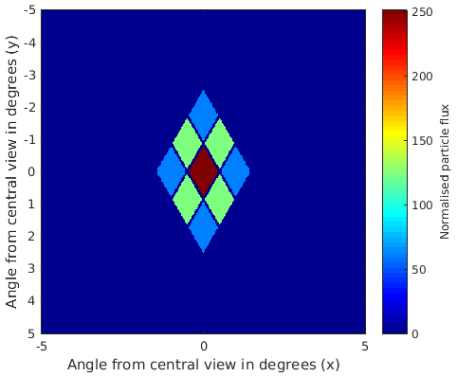
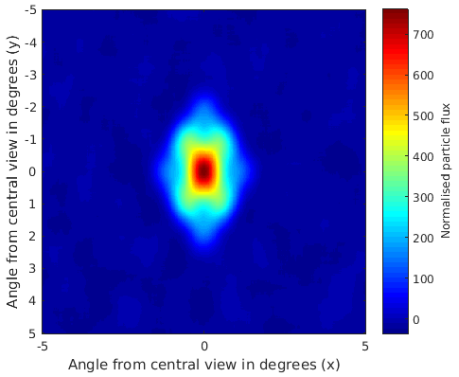


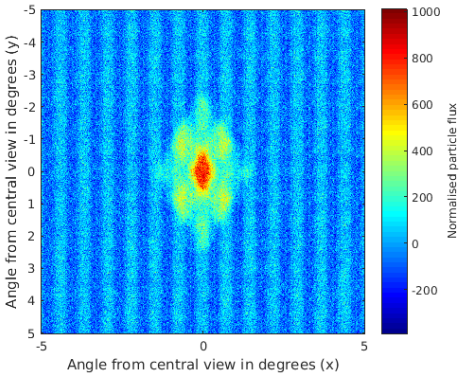
Figure 5.40: The frame resulting from a diamond-based angular distribution described by figure 5.41 below.



(a) Simulated



(b) Reconstructed



(c) Reconstructed

Figure 5.41: The range of angles of 2 keV electrons used in the simulation with their relative intensity, and the reconstructed angular distribution from cross-correlation using the data shown in figure 5.40. (a) is the simulated angular distribution of particles, (b) is the reconstruction using a fully-open mask image and (c) is the higher-resolution but noisier reconstruction using a quarter-open mask image.

However, the behaviour of the strong gradients seen across the detector could not be explained by the MCP-source angular distribution alone. It was suspected that the alignment of the MCP with the mask and CCD setup was poor. Photographic evidence of the chamber setup indicated there could have been a slight angle between the centre line of the detector and mask and the normal to the surface of the MCP source. Numerous simulations were run with either the CCD and mask setup or the MCP source rotated.

Eventually, it was concluded that the source had been set up with a 10° horizontal offset to the right and a 3.5° vertical offset upwards. The detector output from a simulation with the source at this angle using 2 keV electrons is shown in figure 5.42. The differences in intensity across the mask compared to the equivalent data frame in figure 5.20 are likely to be due in a large part to the irregularity of the shapes of the holes in the mask. Although the iterative analysis could have been continued to find a more detailed and realistic angular distribution, it was judged that this reconstruction was adequate for the purposes of demonstration. In fact it was far in advance of the angular resolution the instrument was designed to be capable of measuring.

On a second inspection, the experiment data deconvolution had a peak which matched this to within half of a degree, but in this case no detail could be recovered. The reasons for the difficulty with direct deconvolution could be explained by the use of a source which turned out to be both extended and detailed in both angular and position space and which was angled significantly off-centre.

5.7 Experimental conclusions

The experimental campaign was limited in scope by the available hardware, but allowed some analysis which furthered the understanding of the CCD64 and the sources used. Simulations were used to complement and enhance this understanding.

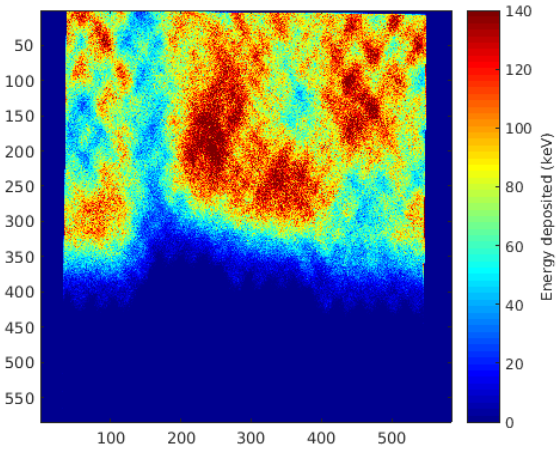


Figure 5.42: The final simulation detector image frame used in the recovery of the properties of the MCP source. This should be compared to figure 5.20 (f).

5.7.1 Assembly

I designed the experimental concept and worked in cooperation with the MSSL design office, workshops and facilities team to produce each component. Clarification of which parts were designed or assembled by the author and others is given here.

Facilities and installation

The vacuum chamber used for the experiments was out of action at the time I needed it. Vitor Botelho found and installed a spare turbo-pump which survived the duration of the experiments and allowed me to use the vacuum chamber.

I installed and aligned all the experiment components and harnessing in the vacuum chamber.

Mask fabrication

I designed and drew the mask in 2-dimensions which was turned into a manufacturing diagram by Craig Theobald in the design office at MSSL. I contacted multiple potential manufacturers to obtain quotes and lead times, which turned out to be prohibitive. On the recommendation of Dhiren Kataria, I then contacted Mark Stanford at the University of Wolverhampton and discussed the manufacturing method and dimensions which he eventually used.

Mechanical interfacing for the components

John Coker in the MSSL design office designed each of the parts to support the mask and CCD64 in the required configuration in consultation with me and in line with my requirements. They were then manufactured in the workshop at MSSL.

Electronics and harnessing

I identified and found the harnesses which could fit to the vacuum chamber. I defined further harnesses to connect these to the hardware I had, including planning how to avoid interference in the sensitive signals. These harnesses were then manufactured by Doug Davies in the electronics workshop at MSSL.

The MCP source was assembled by Richard Darnley in its initial form. I made the modifications necessary to use it at higher voltages for higher electron energies, including diode switchbox modification and harnessing.

Control hardware and software

The PC, electronics and software to run the CCD64 were originally designed by Dave Walton and others and remained in place in the lab from previous tests using the CCD64 with CATS. Dave Walton explained the setup and operation of the power supplies, harnessing and software. All software used to acquire the raw data on the laboratory PC was written by Dave Walton.

I searched for and found the original PC with its associated Inchworm control ISA card to control the Burleigh HV controller which provided the signals to the motors. Both this card and the software were designed by Dave Walton in the early nineties as part of an integrated laboratory control system called the Plasma Automated Chamber Integrated Data Extraction Routine Miscellany (PACIDERM). With the help and explanations of Dave Walton, I was able to set the computer, software, hardware and harnessing up for the motors.

Experiment operation

Vitor Botelho supervised the operation of the vacuum chamber and pump system, and the provision and connection of nitrogen dewars. He and Richard Darnley also moved, installed and removed the two radioactive sources when necessary. I completed all other data collection, logging and analysis.

5.7.2 Outcome of experiments and results

The possible analyses of the mask concept from the experiments using the radioactive sources were limited by the apparent low flux. However, the data taken did allow some interesting conclusions to be drawn.

The thermal behaviour of the CCD64 was quantified to a degree, but was somewhat impaired by image quality, which was probably caused by the erratic nature of the readout electronics.

It was also possible to demonstrate the possibility for the CCD64 to detect electron energy distribution in the 10 keV to 100 keV range, and, to a lesser degree, higher energies. A relatively simple algorithm for identifying particle tracks produced a spectrum which was very well-matched with simulation. The matching of the predicted and modelled energy spectra for all the sources used increases the confidence that the modelled geometry used to represent the CCD64 is close to the actual geometry of the detector, and therefore gives a good indication that simulations involving protons are likely to be valid also for this detector.

It was believed that the MCP source would be able to demonstrate the instrument's performance for a near-parallel source. However, these tests produced an unexpected, complicated and unknown source pattern. Through a long series of iterative simulations, it was possible to produce a possible first order estimate of the source's spatial and angular distribution.

With further investigation, including removing the mask from the setup, it was possible to increase the understanding of the CCD64's response to low energy electrons, around the limit of its detection capabilities.

There were significant limitations in this experimental programme, to a large degree because of the available hardware. The available sources were limited in the scope of what they were able to demonstrate, in particular that they produced electrons only. However proton behaviour in semiconductors is known to be significantly more predictable than electron behaviour.

The exposure time of each frame was limited by the capabilities of the electronics and the exposure time in an experiment run was limited by the cooling time provided by a single dewar of liquid nitrogen. The number of configurations was limited by the fact that to make changes manually, the chamber needed to warm up before it could be opened (to avoid condensation), the changes made, and the pump-down completed, meaning a wait of several days between tests. Without these limitations, the low fluxes of the radioactive sources would have been less of a problem. If an exposure time of a few weeks had been available, it would be expected that a pinhole reconstruction may have been possible.

The CCD64 and associated electronics also introduced further limitations. The CCD itself had some damaged pixels and columns, as well as a radiation damaged section. This was less of a problem for the coded mask technique than for other types of measurement, but still added some difficulties to the analysis. The electronics produced some noise, variable gain, and occasional unusable frames. Again, in some cases this could be compensated for, but the overall quality of what could be tried was limited.

Chapter 6

Applicability to space weather and other missions

The results of the simulations and experiments in the previous chapters show that the coded aperture technique could be expected to be successful in reconstructing highly-directional fluxes.

In reality, the instrument design would be tailored towards the environment it was to be used in and the particles it would be used to detect and analyse.

Environments where the flight of such an instrument might have significant advantages would be ones where a small spacecraft would be economically and scientifically viable and where the environment has a significant population of particles at relevant energies and fluxes. In particular, these would have, or be suspected to have, some directional element.

Small satellites are commonly used in low-Earth orbiting missions, where the low cost of launch allows smaller missions to be worthwhile, but such spacecraft have also been discussed for use in lunar or planetary missions in the future.

Some of the more likely areas for investigation include LEO regions such as the SAA and the auroral regions, and possibly field-aligned particles in the solar wind.

This chapter analyses possible instrument configurations in some of these envi-

ronments.

6.1 Performance of the proof-of-concept instrument as a flight instrument

Although improvements would certainly be possible with improvements in the instrument, especially using a more advanced detector, the instrument as tested still has significant measurement capabilities. It is possible to evaluate the possibility of using the lab-based components as the basis of a flight instrument capable of analysing electrons above 500 eV. It is likely that it could be made to fit within typical resource requirements of a CubeSat or similar satellite.

Although the equipment used for the lab-based tests was very bulky and high-resource, this is in part owing to the fact that the control and amplification electronics were built as part of the test setup for an instrument launched in 2001 and are therefore significantly outdated, and the power supplies and computer hardware and software are similarly outdated and were never designed for flight. Developments in flight hardware would make the electronics for this CCD far more compact and efficient to fly today.

As a demonstration of what improvements in resource requirements might be possible, the lab-based equipment can be compared to the hardware used in the Polar test of the Conceptual And Tiny Spectrometer (PoleCATS) student sounding rocket experiment which used a very similar CCD64 for low-energy electron detection in Lee et al. [2013].

Figures 6.1 and 6.2 show the lab-based setup used outside the chamber for CCD control. The PC was used for initialising and running the CCD control signal generation hardware, with the possibility of adjusting exposure time, and also acquiring and storing image data. The Power Supply Units (PSUs) provided all five necessary voltages for the control electronics, seen in the upper left of the image, to run the CCD. The final electronics board, for signal amplification, was located in the

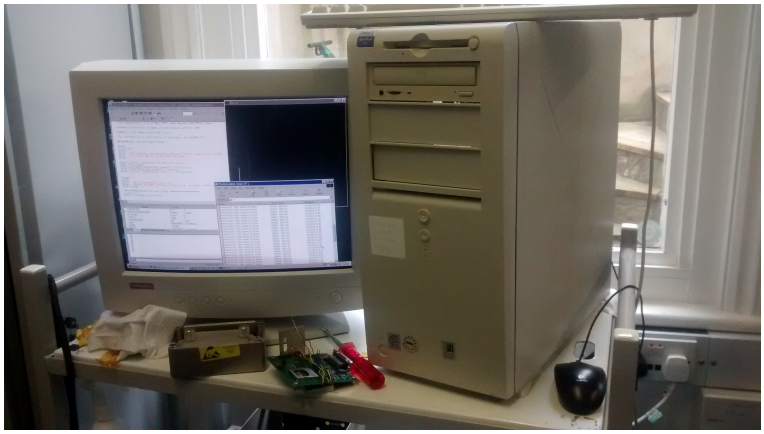


Figure 6.1: The PC used for controlling the CCD operation and readout

vacuum chamber.

Figure 6.3 shows the updated electronics used on the sounding rocket platform in 2013. The logic chips the older system was built around have been replaced by Integrated Circuits (ICs) designed to run CCDs. A PIC24 microcontroller and an AD9920A CCD front end from Analog Devices were capable of running and reading out the CCD with approximately the same readout time as the lab-based electronics. This board supplied CCD control signals and connected necessary voltages to the CCD. It also read out and processed analogue data from the CCD.

Another PC/104 board with lower population density provided all voltages required for the experiment systems including CCD voltages. Many of these voltages would be produced on a typical satellite payload, and it is therefore likely that such a CCD could be expected to be controlled and read out using only one PC/104 PCB.

However, a CCD64 would also encounter some issues on a CubeSat platform, in particular from light sensitivity and thermal noise. Limited mitigation of these issues would be possible, for example by blackening all surrounding components to avoid stray light and taking measurements when the Sun and the Earth are outside

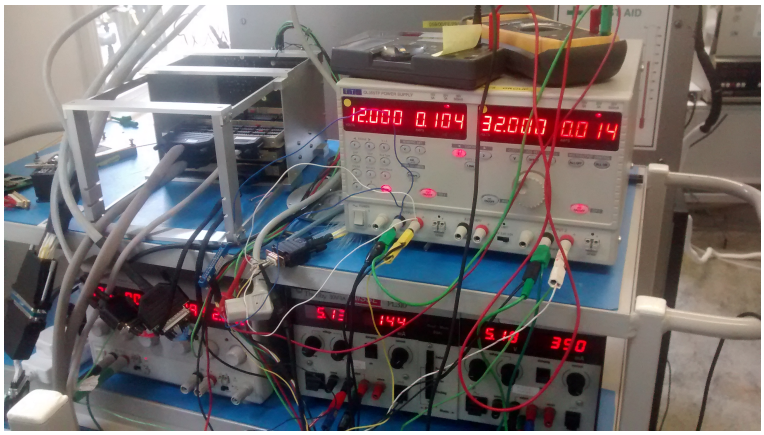


Figure 6.2: Power supplies and digital signal generation electronics

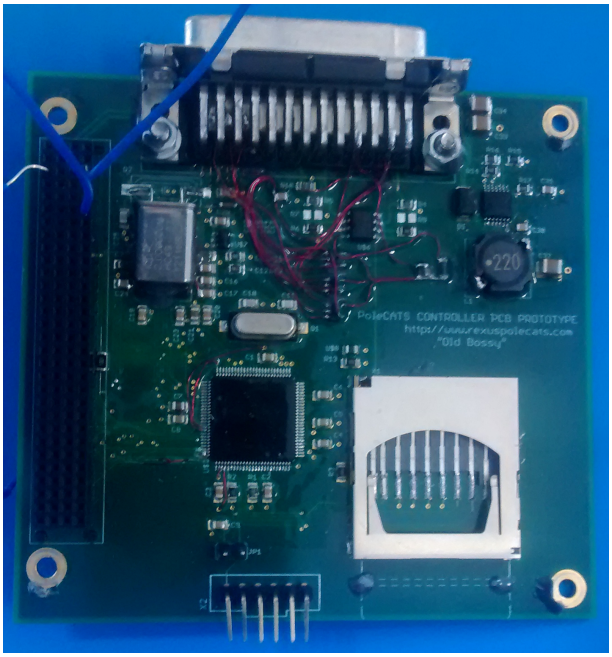


Figure 6.3: PC/104 form factor PCB providing CCD signals, readout, processing and data storage

the FOV.

Since the temperature of a CubeSat might be expected to vary between as much as −30 °C to 20 °C through an orbit, for the nightside of an orbit, the detector might expect to function fairly well. However, there would be very little chance of it working well on the dayside. The temperature, sunlight and Earth's albedo would all be likely to cause significant problems for this detector. It would be prudent to blacken the mask and surrounding internal surfaces near the detector.

For a flight instrument, the most suitable alternative detector is likely to be one of the Medipix family. These detectors are very recent and advanced technology and are rapidly gaining flight experience, as well as a large amount of experience in other fields.

Medipix and Timepix detectors have the strong advantages afforded by the flexibility in the mode of operation. This flexibility offers the possibility of energy discrimination, or of analysing tracks of individual particles to recover, for example, the particle species. The relative weakness of the Medipix family of detectors when compared to the CCD64 is the lower spatial resolution. However, the CCD64 pixel size was unnecessarily small when compared to the scale of the instrument, so the Medipix design is more suitable and therefore requires slightly less computation time and power to process. Such a change in detector would require updated geometry and data handling as well.

6.2 Simulation of the instrument in a predicted LEO environment

A simple model was built in Geant4 to perform omnidirectional simulations of possible radiation environments taking into account the satellite housing around a proposed instrument.

This model was built around a detector with a Medipix geometry - a surface area of 14.08 mm×14.08 mm with a dead layer and active region thickness that can be

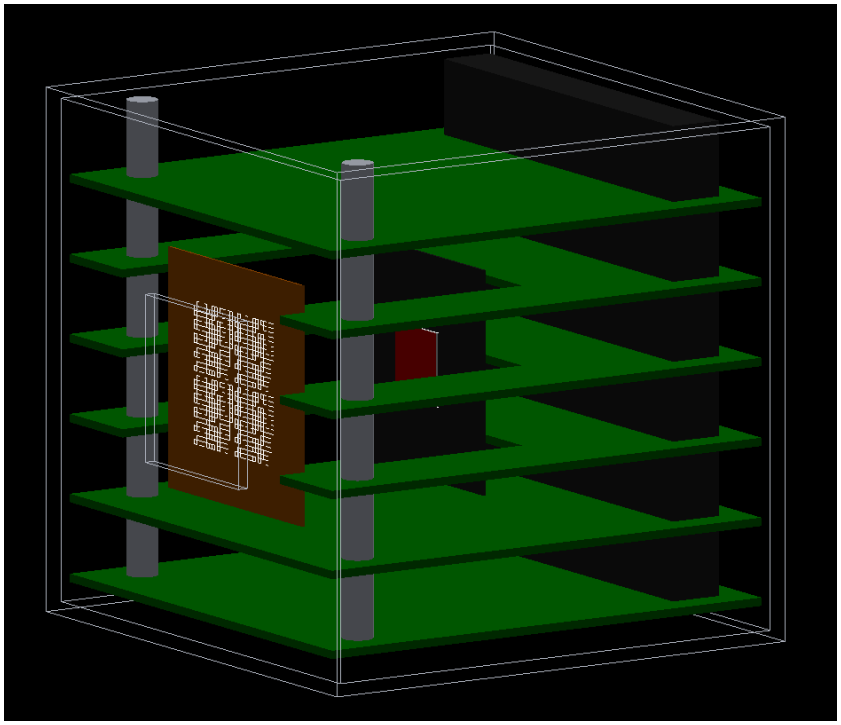


Figure 6.4: A 1U CubeSat carrying a coded aperture instrument imagined with six PC/104 circuit boards.

adjusted. In these simulations the dead layer thickness can be chosen for particles of interest in a particular mission.

The electronics for detector control, data handling and processing, and provision for a bias voltage for the detector also needed to be fitted into this geometry.

In this environment, a model of the detector and mask placement can be used, set into the side of a standard CubeSat chassis. A model of this in Geant4 is shown in figure 6.4. In this case, the chassis is shown as a wireframe, and a possible structure of six boards based on the PC/104 form factor are fitted around the instrument, with cutouts to accommodate it where needed.

The complete model used for simulations is shown in figure 6.5, with a complete

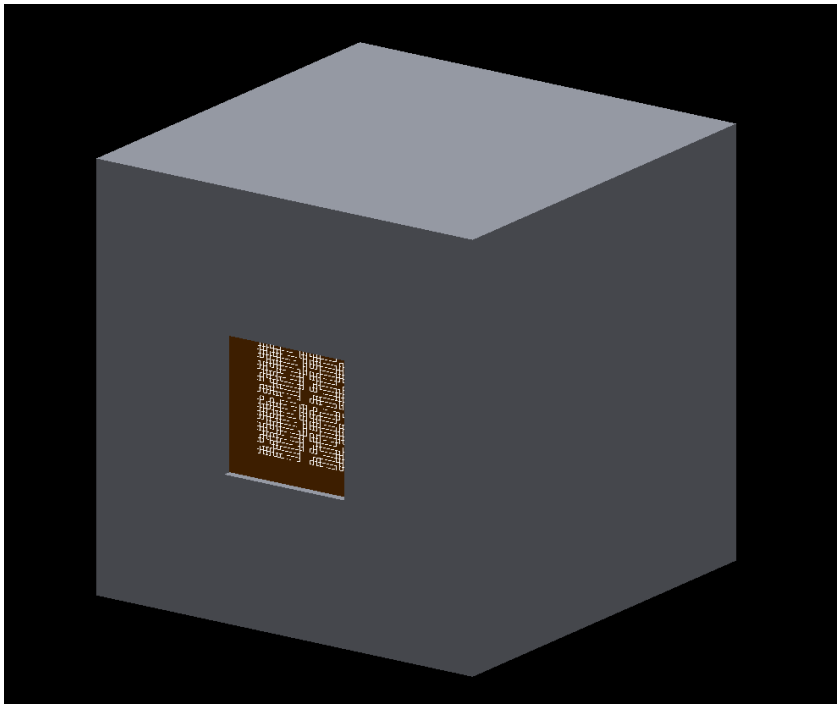


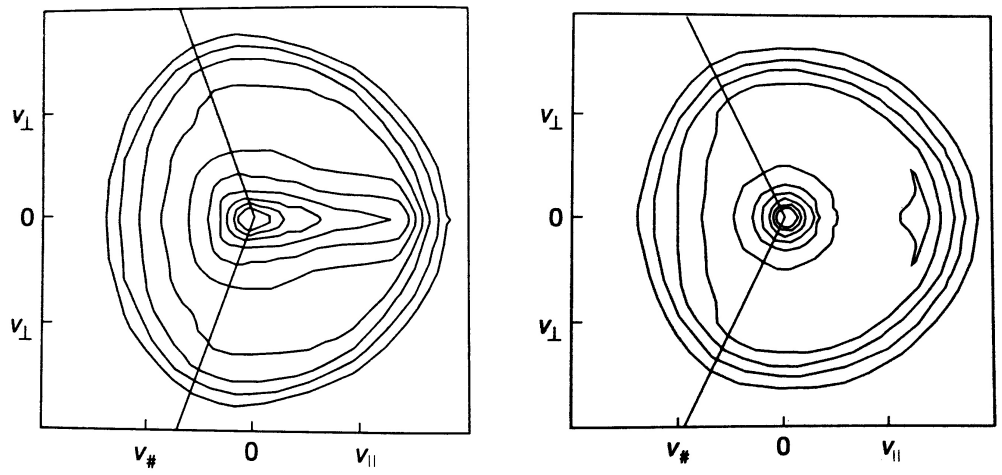
Figure 6.5: A 1U CubeSat-sized aluminium box of 10 cm cubed with a plate thickness of 2 mm

chassis covering each side with 2 mm of aluminium. This provides adequate shielding for the instrument and electronics in a magnetospheric radiation environment.

6.2.1 Measurements in the auroral regions

In a low-Earth orbit environment, one of the key measurements is at the auroral regions where field lines converge. Here, field-aligned electrons penetrate into the upper ionosphere and may interact with the outer reaches of the atmosphere.

These regions between 300 km to 700 km altitude have been sampled both by rockets and by low-orbiting spacecraft. Sounding rocket measurements into auroras, for example those by McFadden et al. [1987] which used an electrostatic analyser,



(a) (b)

Figure 6.6: Angular and velocity distribution measured by McFadden et al. [1987] and redrawn by Bryant [1998].

have observed an approximately isotropic background and a more strongly directional component aligned along the magnetic field lines. An example of plots of these distributions are shown in figure 6.6. The directional component may be a spectrum or a narrower peak of energies, but in the set of results from McFadden et al. [1987], this was up to around 10 keV.

Depending on the operations of a CubeSat it may have some rotation. In this case, it is assumed the detector is able to take a frame of data at a maximum exposure time of 1 s without rotating more than 1°.

For an instrument aimed at measuring these environments, a detector geometry with a 100 nm entrance window was chosen to be simulated, which would be expected to detect electrons above approximately 2 keV.

The environment simulation was an omnidirectional electron flux with a power spectrum of $E^{-1.73}$, and a 5° wide flux with an energy centred around 10 keV with

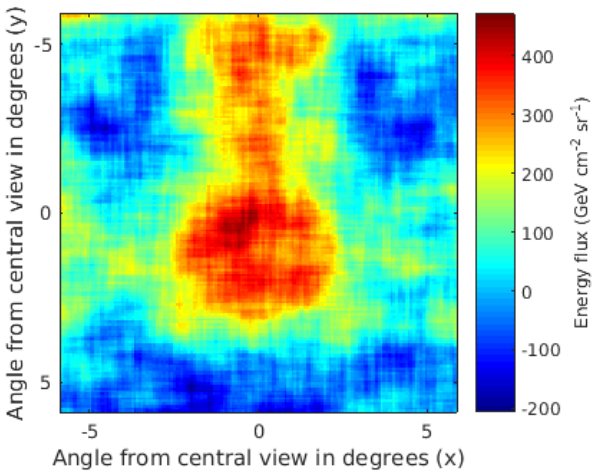


Figure 6.7: A reconstructed field-aligned flux of 10 keV electrons from a simulated 1 s exposure time.

an intensity two orders of magnitude above the 10 keV flux in the rest of the angular distribution. The aligned flux was a small arbitrary angle away from the normal to the face of the spacecraft on which the instrument was mounted.

The result of a simulated 1 s exposure is shown in figure 6.7. This is with an arbitrary detector without spectral response, so recovering the energy of the incident particles in the flux was not possible; only the total deposited energy from the flux could be recovered.

For slightly more distributed sources, it may become more effective to choose a coarser aperture shape, such as an 8×8 PBA, so that the shadow patterns produced on the detector are clearer.

6.3 Consideration of the instrument in other environments

Other environments which are likely to have strongly-aligned particle fluxes include the SAA, which has few measurements at low altitude and relatively poor detail in the AE-9 and AP-9 models of <100 keV electrons. There are also other such environments outside the Earth's magnetosphere, where CubeSats are not currently flown. Measurements could be taken of the electron strahl, which is described by Pierrard et al. [2001] and Anderson et al. [2012], given the right detector. Currently top hat detectors take very successful measurements at low energies, but a pinhole or coded mask instrument could be manufactured on a much smaller scale and therefore fly on a wider range of missions.

Both Saturn and Jupiter have been observed to have very bright auroras with very high energy fluxes. Jupiter's high-energy electron radiation belts have been measured to contain electrons at energies up to 50 MeV. Blocking these particles with a mask or shielding would require nearly 1 cm of tungsten, and this would be likely to produce bremsstrahlung without careful shielding design.

6.4 Development of the concept into a flight instrument

If the instrument was to be used for flight a number of parameters and factors need to be considered, including the choice of the detector and geometry, but also data processing, interfacing to the satellite and the thermal environment.

6.4.1 Mechanical considerations

An instrument built into a CubeSat would use the edges of a hole in the face of the spacecraft structure as a baffle and additional shielding against incident particles

outside its field of view.

The instrument placement in a CubeSat would depend on the attitude control intended for the satellite. A satellite with a slow rotation about an axis in the direction of travel would be able to sweep between zenith, nadir and the angles in between. A 3-axis stabilised satellite would offer the possibility of constantly monitoring a given direction of interest. Ideally, more than one instrument would be mounted on the spacecraft, for example with nadir and zenith pointing.

The placement also depends on the type of detector, and the need for protection from light sources. A detector that is not sensitive to visible light would be simpler to position. CMOS detectors might be an interesting detector candidate due to the simplicity of the required electronics, as all analogue processing is done within the detector.

Thermally, the detector is likely to be reasonably safe from the strongest temperature fluctuations, since it is recessed from the CubeSat surface and somewhat shaded. The CCD64 had the disadvantage of needing to be cooled to below approximately $-40\,^{\circ}\mathrm{C}$ to produce the clearest images. The cooling requirements of a Medipix depend on the detector used. Some may require temperatures as low as $-50\,^{\circ}\mathrm{C}$, while others can operate at room temperature - matching the capability to the requirements of the instrument and platform is a key decision.

High radiation environments pose a risk of radiation damage to the detector. For the case of the Medipix, the sensitivity to radiation damage depends on the structure of the bulk of the detector. If a background without incident particles can be measured or pieced together, to establish the dark current produced by radiation damage, the damaging effect can be minimised, as was done with the radiation-damaged section of the CCD64 in the laboratory tests. At the end of the detector's life, radiation damage could become severe enough that the detector is either unusable or it must be annealed. For a CubeSat mission, annealing is unlikely to be used, depending on the mission's planned lifetime.

The required electronics need to be miniaturised. Using modern electronics and

standard PCB sizes this could be accomplished using around half of a CubeSat unit for the whole instrument while having powerful enough signal processing to be able to handle the generated images. It could even be possible to reduce the instrument to an eighth of a CubeSat unit, 5 cm×5 cm×5 cm, by optimising the PCB placement on the outer walls of the cavity between detector and mask. The available space should be sufficient to place all required electronics.

A sounding rocket platform would also be suitable for this type of miniaturised standalone instrument. It would be possible to position 8 separate instruments around the circumference of a medium-altitude sounding rocket and get 360° by 56° coverage with overlap between adjacent detectors. A rocket moves quickly through the features of interest and in this case a cadence at a maximum exposure time of 1 s would be needed and probably significantly faster unless the rocket was de-spun.

6.4.2 Operations

On a spinning CubeSat the choice of measurement cadence depends on the surrounding particle flux and the motion and rotation speed of the satellite.

If the instrument cannot sustain continuous operation either due to constraints in the power budget or limited telemetry then measurements only within the regions of interest, for example around the auroral regions or the SAA, would be possible using either timed commands or on-board trajectory information.

The available downlink of the data generated by the instrument might be limited. To utilise the available downlink it would be possible to either fully process the frames on board and reduce their resolution, or to compress the raw images. As on-board data storage is generally not an issue due to the availability of modern high-density storage, a snapshot or reduced data could be downlinked for assessment, and then the required data could be downloaded.

It is also worth noting that when the environment is too low flux or dynamic or unpredictable for the instrument to work at high resolution as intended, the aperture size as a whole can still provide useful information about the flux levels over the

solid angle subtended by the FOV.

6.4.3 Characterisation and calibration

A full characterisation and calibration of a coded aperture instrument should include an understanding of the instrument response to particle type, particle energy, and incident angle using well-known and previously calibrated sources.

As a step towards increasing the Technology Readiness Level (TRL) of the instrument, it could be demonstrated on a sounding rocket platform in the auroral regions.

Chapter 7

Conclusions

The concept's demonstrated and simulated performance shows that it is worth further development for environments with low flux that is, or is suspected to be, strongly directional.

A summary is given here of the stages of design, simulation and development of the instrument to this point, and possible avenues of future research of the concept.

7.1 Summary of the instrument development

A review of historical and current *in-situ* particle analysers for space radiation was made, with a particular focus on the new generation of miniaturised designs that are now feasible to manufacture and possible to fly. These have found more applications with the increase in number and reliability of small satellite, including CubeSat, missions. The aim was to produce a novel miniaturised low-resource plasma analyser on this scale for space weather research.

Current miniaturised particle analysers are hampered by low geometric factors and therefore low sensitivity. As an alternative, the coded aperture technique, developed in the 1960s for X-ray astronomy and applied to several particle radiation sources in the last decade, offers a higher geometric factor with a high angular resolution, and some potential to identify the incident particles.

The use of this technique comes together with need to improve space weather measurements and predictions. The size and simplicity of such instruments allows them to take advantage of the CubeSat swarm and constellation mission possibilities and provide high-accuracy simultaneous multi-point measurements

7.1.1 Design of the proof-of-concept instrument

The starting point of the design was based around the scale of commercially-available position sensitive detectors, with the aim of fitting an entire instrument with its supporting electronics into approximately a CubeSat form. The choices of the mask shape and scale, along with the necessary thickness and material and the eventual deconvolution algorithms were based around the history and theory of the coded aperture technique and particle transport.

Extensive simulations have been performed to identify the optimum configuration and to determine key parameters of the design. Based on these, a proof-of-concept instrument was assembled for testing in the laboratory.

7.1.2 Performance of the proof-of-concept instrument

A proof-of-concept version of the instrument was fabricated and then tested. A specialised ion-implanted CCD64 with an exceptionally thin entrance window was used as the detector and the mask, scaled to match the CCD, was manufactured from a tungsten-copper pseudo-alloy using additive manufacturing. The setup was tested in a vacuum chamber with three different types of electron source.

Although each of these electron sources presented some surprises in their behaviour, important lessons were learned. The two radioactive sources exhibited very low activity levels which were not sufficient to determine the coded aperture concept, but demonstrated the ability of this system to measure the energy spectrum of incident radiation. The third source, an MCP-based electron source, was then used to verify the coded mask. Even though neither the source nor detector were previously characterised, a good understanding of both the instrument and the

source's angular and spatial distributions, and the behaviour over different acceleration voltages was gained through co-simulation using Geant4.

The performance of the proof-of-concept instrument in the laboratory setup was generally as expected, but further characterisation was not possible due to time and hardware constraints.

The testing was limited to electron sources only, since no proton or ion sources were available, so any predicted response for other particles relies on simulations only. In addition, it was not possible to demonstrate the reconstruction of the distribution of a known source with the measurements that were taken, for reasons of low flux and high complexity. However, the experiment data were used to diagnose some of the characteristics of the possible features of the MCP source, and the further simulations produced a similar response to some of the experimental measurements.

7.1.3 Simulation of a possible flight instrument configuration

Simulations were run of an omnidirectional auroral environment surrounding a model of a CubeSat design in Geant4. Where a source is very widely distributed, reconstruction is significantly more difficult. However, gradients from one side of the FOV to the other can in general be recovered.

7.2 Main results

The concept was developed based on previous instruments and inspired by the development of the CubeSat family of satellite platforms, and was subjected to testing in simulation and experiment.

Simulations developed in the Geant4 toolkit were used to predict the possible capabilities of the instrument and refine the design of the geometry. It was demonstrated that the proposed geometry was capable of distinguishing two parallel beams with 2° between them, even in environments with high background flux or high de-

tector noise. Further investigation showed the, somewhat reduced, potential for measuring more distributed sources.

The SNR that could be expected depends on the distribution of the source as well as the geometry of the instrument. It was demonstrated that significantly better SNRs was possible with a mathematically-ideal mask design, and with a repeated mask shape, which also allowed a fully-coded field of view of 20° and a partially-coded field of view of 56°. For this final geometry, numerous simulations of angular distributions of electrons were simulated, which produced SNRs of between 17 and 464, largely depending on the complexity of the distribution.

It is clear from these simulations that the particular strength of the coded aperture technique for particle measurements is when the angular distribution of a source is narrow and well-defined in structure. The ability to recover more distributed sources is significantly weaker, but possible.

A proof-of-concept version of the instrument was fabricated and then tested. A specialised CCD with an exceptionally thin entrance window was used as the detector and a mask was manufactured from a tungsten-copper pseudo-alloy. The setup was tested in a vacuum chamber with three different types of electron source. Although each of these sources presented some surprises in their behaviour, important lessons were learned. The two radioactive sources exhibited very low activity levels which were not sufficient to determine the coded aperture concept, but demonstrated the ability of this system to measure the energy spectrum of incident radiation. The third source, an MCP-based electron source, was then used to verify the coded mask. Even though neither the source nor detector were previously characterised, a good understanding of both the instrument and the source's angular and spatial distributions, and the behaviour over different acceleration voltages was gained through co-simulation using Geant4.

No proton or ion sources were available, so the testing was limited to electron sources only. The predicted responses for other particles rely on simulations only. In addition, it was not possible to demonstrate the reconstruction of the distribution

of a known source with the measurements that were taken, for reasons of low flux and high complexity. However, the experiment data were used to diagnose some of the characteristics of the possible features of the MCP source, and the further simulations produced a similar response to some of the experimental measurements.

This design uses heritage from space-based X-ray telescopes and particle detectors, and ground-based instrumentation including the gamma-source mapping technology described by Gmar et al. [2011] and Lemaire et al. [2014]. It demonstrates the first investigation into the use of the coded aperture technique for the detection of charged particles in space.

7.3 Possible future development of the instrument

If the laboratory-based version of the instrument were to be flown, at a minimum it would need to be able to withstand the changing radiation and thermal environments. The electronics used for the lab based tests would need to be miniaturised into a small form factor, which would be feasible using modern electronics. Radiation-hardened components would need to be used for the electronics, to a level suitable for the flight environment, and the radiation hardness of the detector should be established. It would also be likely to be cost-effective to blacken the mask and surrounding components to eliminate stray light reflections.

In addition, the most interesting fields of research for instrument development might be using a more versatile detector and associated deconvolution algorithm, manufacturing additional structure in the mask, and the ability to move or change the design of the mask. A brief description of these proposed developments is given here.

7.3.1 Detector development

The most obvious, and desirable, improvement to the instrument concept as developed in the laboratory would be the substitution of a more suitable or capable

detector for the CCD64.

The Medipix range of sensors provides many more options of readout schemes, and the flexibility to choose the geometry and material of the detector used on the chip.

Should a Medipix detector be used, there are likely to be some limited possibilities to identify both individual particles and their incident angle with a certain degree of accuracy, depending on software and physical limitations. The use of a mask such as the ones described here would be able to increase the accuracy of this angular information. Investigation into such a setup and the associated data handling algorithms is beyond the scope of this thesis, since it was not possible to acquire a Medipix, but would be of great interest if or when one could be acquired.

The Timepix chip is capable of time-over-threshold measurements, which gives an indication of the energy deposited in a pixel during a particle interaction. Using these measurements, it should be possible to recover the incident energy spectrum and an estimate of angular distribution for each energy in it. Using the advanced features of a detector in the Medipix family, it would also be possible to determine the angular distribution for a wide range of energy bins in a single detector.

The light sensitivity of the detector used is a key limitation to the capabilities of an instrument. The thinness of the entrance window would need to be set to meet the requirement for the lower bound for the planned energy measurements. This thickness would also have a large impact on the sensitivity to light of the instrument. Balancing the requirements of the detector with the exposure of the instrument or orientation of the spacecraft would be critical to acquire useful measurements and avoid damage to the detector.

7.3.2 Improvements of the data analysis

Adding a Medipix detector to coded mask technique as described above would allow the development of new algorithms; the hits could be stored and binned into frames after the data was collected for processing or downlinking.

A Medipix-specific deconvolution algorithm would include both the rough prediction of the energy and angle of each incident particle and the additional information from the knowledge that the particle must have passed through one of the open holes. This combination of inputs to the algorithm could potentially provide high quality information of both total flux within the instrument FOV and continuously-increasing statistical information on high-resolution fluxes. For example, an instrument on a rotating platform could use the time and position information in each hit to reconstruct an angular distribution without the ambiguity associated with a longer exposure time.

The gamma source imaging technologies described by Gmar et al. [2011] and Lemaire et al. [2014] may have some parts in common with this type of analysis software, but these would need significant modification to be optimised for the higher-flux, more distributed angular sources.

Beyond choosing the most effective deconvolution algorithm to take into account properties of the expected background noise and other factors, a flight data handling plan will be highly dependent on the telemetry requirements and processing power available on a platform. Generally, it is considered preferable to have raw image data transmitted to the ground station so that all stages of processing can be observed and checked on the ground. In the case of this instrument, it would be advantageous to recover data from before the deconvolution step, since all details about any individual particles are lost in deconvolution.

In a CubeSat, telemetry data-rates may not be high enough to be able to recover all the raw data, so the best choice must be made between data processing, compression, and selection of the most relevant frames. The analysis of how to implement these strategies in different situations must be optimised and included in the flight software.

7.3.3 Techniques

If the resource budget of a platform allowed it, an additional component in the setup would allow a further degree of customisation for a planned mission.

Mask development

A low-resource addition to the mask concept would be modifying the shape of the holes. This may be useful in situations where the instrument and by extension the mask holes are small, but the mask must be thick to block high energy particles. A design using holes at various angles to allow flux from a wider range of angles was proposed by Hong et al. [2004]. However, such a development would make the manufacturing and the image deconvolution significantly more involved. The sintering fabrication technique of the mask used in the experiments described in chapter 5 would be ideal for this type of design. The behaviour of the shadow pattern of a more complex and three-dimensional mask shape with a granular structure from manufacturing could be identified optically using a light source and a light-sensitive position sensitive detector, with no need to run the experiment under vacuum.

A rotating mask

The concept of a rotating mask requires only one moving component which could be controlled by a single motor. In particular, the hexagonal design described by Cook et al. [1983], which could be used with a modified Medipix detector design, is particularly elegant, and is calculated to nearly eliminate artifacts arising from irregularities or damage in the detector. However this would require a 50% open mask design that could be manufactured and self-supporting - this would probably be achieved most easily by adding narrow connections to hold the mask elements together.

A related construction used in a different concept might have shutter that would close off all or part of the mask when necessary. When the flux detected is high and the coded mask is not necessary, the open area of the mask could be reduced to the

size of one pinhole. This would reduce the instrument to a far simpler design with a more accurate angular reconstruction. This would however be at the expense of reducing the FOV - to 20° in the case of the instrument simulated here.

Another way of implementing this, if space and mass constraints allowed it, would be a mask wheel which could be used to select the aperture shape. For example, it might be possible to choose between a pinhole, a URA or a repeated URA depending on the detected flux or mission requirements.

7.3.4 Further characterisation and flight

A calibrated particle source with a more even spatial distribution should be used to more fully characterise a prototype or flight instrument. In addition, a full instrument characterisation would cover not just incident electrons, but also protons and ions. A radioactive alpha source with sufficient activity level would be enough to demonstrate the capability of the instrument with alpha particles.

The proposed design for the instrument could be demonstrated in an environment with auroral electrons on a sounding rocket platform with an altitude of at least 300 km with minimal additional development required, to raise the TRL of the technique.

This instrument would be ideal for swarm missions, being small, simple, capable, low-resource and possibly low-cost, depending on the detector.

7.4 Concluding remarks

This novel method for space-based in-situ particle analysis has been developed from a starting point of its previous uses as a large-scale high-angular-resolution instrument in X-ray and gamma ray astronomy into a highly-miniaturised low-resource energetic particle instrument which could be used in a range of particle environments.

As the capabilities of CubeSats and other small satellite platforms increase and they become more widely-used as full scientific missions, miniaturised instruments

such as this one will significantly extend the range of magnetospheric scales over which models can be improved and verified.

Acknowledgements

This PhD was funded by STFC. Additionally, I would like to mention and acknowledge the assistance of:

- Alan Smith and Dhiren Kataria, for the opportunity, their supervision and their support.
- Dave Walton, who was always willing to help fight the CCD.
- Vitor Botelho, who went to so much effort to make sure my experiments happened.
- Craig Theobald, John Coker, Doug Davies and Graham Willis for their help with the design and manufacturing.
- Mark Stanford, at the University of Wolverhampton, for manufacturing the mask.
- Rob Bedington and the PoleCATS team, for showing me the interest and the relevance in detectors.
- Tom Nordheim, for proof reading and advice.
- All my colleagues, friends and family for their patience and help.

References

Abel, B. and R. M. Thorne

1999. Modeling energetic electron precipitation near the South Atlantic anomaly. *Journal of Geophysical Research*, 104(A4):7037–7044. 54

Ables, J.

1968. Fourier transform photography: a new method for X-ray astronomy. *Proceedings of the Astronomical Society of Australia*, 1(4):172–173. 145

Alfvén, H.

1945. Magneto-hydrodynamic waves and sunspots. *Monthly Notices of the Royal Astronomical Society*, 105:3–16. 42

Althouse, W., W. Cook, A. Cummings, M. Finger, D. Palmer, T. Prince, S. Schindler, C. Starr, and E. Stone

1987. First flight of a new balloon-borne gamma-ray imagine telescope. *Proceedings of the 20th International Cosmic Ray Conference Moscow*, 1:84. 158

Amelio, G. F., Tompsett M. F., and Smith G. E.

1970. Experimental verification of the charge coupled device concept. *Bell System Technical Journal*, 49(4):593–600. 113

Amendolia, S. R., E. Bertolucci, M. G. Bisogni, U. Bottigli, A. Ceccopieri, M. A. Ciocci, M. Conti, P. Delogu, M. E. Fantacci, P. Maestro, V. Marzulli, E. Pernigotti, N. Romeo, V. Rosso, P. Rosso, A. Stefanini, and S. Stumbo
1999. MEDIPIX: A VLSI chip for a GaAs pixel detector for digital radiology. *Nuclear Instruments and Methods in Physics Research A*, 422(1-3):201–205. 129

Anderson, B. R., R. M. Skoug, J. T. Steinberg, and D. J. McComas
2012. Variability of the solar wind suprathermal electron strahl. *Journal of Geophysical Research: Space Physics*, 117(4):A04107. 48, 304

Anderson, K., L. Chase, R. Lin, J. McCoy, and R. McGuire
1974. Lunar particle shadows and boundary layer experiment: Plasma and energetic particles on the Apollo 15 and 16 subsatellites. Technical report, California University. 14, 83

Aschwanden, M.
2006. *Physics of the solar corona: an introduction with problems and solutions*. Springer Science & Business Media. 43, 46

Auster, H. U., K. H. Glassmeier, W. Magnes, O. Aydogar, W. Baumjohann, D. Constantinescu, D. Fischer, K. H. Fornacon, E. Georgescu, P. Harvey, O. Hillenmaier, R. Kroth, M. Ludlam, Y. Narita, R. Nakamura, K. Okrafka, F. Plaschke, I. Richter, H. Schwarzl, B. Stoll, a. Valavanoglou, and M. Wiedemann
2008. The THEMIS fluxgate magnetometer. *Space Science Reviews*, 141(1-4):235–264. 70

Baker, D. N., A. N. Jaynes, V. C. Hoxie, R. M. Thorne, J. C. Foster, X. Li, J. F. Fennell, J. R. Wygant, S. G. Kanekal, P. J. Erickson, W. Kurth, W. Li, Q. Schiller, L. Blum, D. M. Malaspina, A. Gerrard, and L. J. Lanzerotti
2014. An impenetrable barrier to ultrarelativistic electrons in the Van Allen radiation belts. *Nature*, 515(7528):531–4. 54

Baker, D. N., S. G. Kanekal, V. C. Hoxie, S. Batiste, M. Bolton, X. Li, S. R. Elkington, S. Monk, R. Reukauf, S. Steg, J. Westfall, C. Belting, B. Bolton, D. Braun, B. Cervelli, K. Hubbell, M. Kien, S. Knappmiller, S. Wade, B. Lamprecht, K. Stevens, J. Wallace, A. Yehle, H. E. Spence, and R. Friedel
2013. A long-lived relativistic electron storage ring embedded in Earth's outer Van Allen belt. *Science*, 340(6129):186–190. 54

Banks, P. M., C. R. Chappell, and A. F. Nagy
1974. A new model for the interaction of auroral electrons with the atmosphere - Spectral degradation, backscatter, optical emission, and ionization. *Journal of Geophysical Research*, 79(10):1459–1470. 57, 58

Becquerel, H.
1900. Contribution a l'étude du rayonnement du radium. *Journal of Theoretical and Applied Physics*, 9(1):190–199. 78

Bedington, R.
2012. *A prototype Cylindrical And Tiny Spectrometer for the rapid energy analysis of space plasmas*. PhD thesis, UCL (University College London). 123

Bedington, R., D. Kataria, and D. Walton
2012. Using a CCD for the direct detection of electrons in a low energy space plasma spectrometer. *Journal of Instrumentation*, 7(01):C01079–C01079. 89, 223, 225, 252, 263, 265, 274

Belehaki, A., I. Stanislawska, and J. Lilensten
2009. An overview of ionosphere-thermosphere models available for space weather purposes. *Space Science Reviews*, 147(3-4):271–313. 63

Benkhoff, J., J. van Casteren, H. Hayakawa, M. Fujimoto, H. Laakso, M. Novara, P. Ferri, H. R. Middleton, and R. Ziethe
2010. BepiColombo-Comprehensive exploration of Mercury: Mission overview and science goals. *Planetary and Space Science*, 58(1-2):2–20. 66

Bethe, H.
1930. Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie. *Annalen der Physik*, 397(3):325–400. 102

Bethe, H.
1932. Bremsformel für Elektronen relativistischer Geschwindigkeit. *Zeitschrift für Physik*, 76(5-6):293–299. 106

Bethe, H. and R. Bacher
1936. Stationary States of Nuclei. *Reviews of Modern Physics*, 8:82–229. 275

Bömer, L. and M. Antweiler
1990. Periodic Complementary Binary Sequences. *IEEE Transactions on Information Theory*, 36(6):1487–1494. 152

Boyle, W. S. and G. E. Smith
1970. Charge Coupled Semiconductor Devices. *Bell System Technical Journal*, 49(4):587–593. 113

Braga, J., F. D'Amico, and T. Villela
1994. Development of the hard X-ray imaging experiment timax: Laboratory images and first balloon flight. *Experimental Astronomy*, 5(3-4):269–278. 157

Braga, J., T. Villela, U. B. Jayanthi, F. D'Amico, and J. A. Neri
1991. A new mask-antimask coded-aperture telescope for hard X-ray astronomy. *Experimental Astronomy*, 2(2):101–113. 155, 157

Brandt, S., M. Hernanz, L. Alvarez, A. Argan, B. Artigues, P. Azzarello, D. Barret, E. Bozzo, C. Budtz-Jørgensen, R. Campana, A. Cros, E. del Monte, I. Donnarumma, Y. Evangelista, M. Feroci, J. L. Galvez Sanchez, D. Götz, F. Hansen, J. W. den Herder, R. Hudec, J. Huovelin, D. Karelin, S. Korpela, N. Lund, M. Michalska, P. Olsen, P. Orleanski, S. Pedersen, M. Pohl, A. Rachevski, A. Santangelo, S. Schanne, C. Schmid, S. Suchy, C. Tenzer, A. Vacchi, D. Walton, J. Wilms, G. Zampa, N. Zampa, J. in't Zand, S. Zane, A. Zdziarski, and F. Zwart
2014. The design of the wide field monitor for the LOFT mission. 9144:91442V. 159

Bryant, D.
1998. *Electron acceleration in the aurora and beyond*. CRC Press. 20, 302

Busboom, A., H. Elders-Boll, and H. D. Schotten
1998. Uniformly redundant arrays. *Experimental Astronomy*, 8(2):97–123. 15, 150, 151, 152

Busboom, A., H. D. Schotten, and H. Elders-Boll
1997. Coded aperture imaging with multiple measurements. *Journal of the Optical Society of America A*, 14(5):1058–1065. 158

Byard, K.
1992. On self-supporting coded aperture arrays. *Nuclear Instruments and Methods in Physics Research A*, 322:97–100. 149, 158

Byard, K.
2014. Fast decoding algorithms for coded aperture systems. *Nuclear Instruments and Methods in Physics Research A*, 754:36–41. 156

Calabro, D. and J. K. Wolf
1968. On the synthesis of two-dimensional arrays with desirable correlation properties. *Information and Control*, 11(5-6):537–560. 150

Carnelli, I., A. Galvez, K. Mellab, and M. Kueppers
. Asteroid Impact Mission: a unique opportunity to demonstrate planetary defense. In *4th IAA Planetary Defense Conference*. 74

Caroli, E., J. B. B. Stephen, G. Cocco, L. Natalucci, and A. Spizzichino
1987. Coded Aperture Imaging in X- and Gamma-Ray Astronomy. *Space Science Reviews*, 45(3-4):349–403. 145

Carrington, R.
1859. Description of a Singular Appearance in the Sun. *Monthly Notices of the Royal Astronomical Society*, 20:13–15. 61

Cohen, B., R. Sellar, R. Staehle, N. Toomarian, and D. Paige
2013. Lunar FlashLight: mapping lunar surface volatiles using a CubeSat. In *Annual Meeting of the Lunar Exploration Analysis Group*. 74

Cohen, B. L.
1971. *Concepts of Nuclear Physics*. McGraw-Hill. 274

Collinson, G. A. and D. O. Kataria
2010. On variable geometric factor systems for top-hat electrostatic space plasma analyzers. *Measurement Science and Technology*, 21(10):105903. 80

Cook, W. R., M. Finger, T. A. Prince, and E. C. Stone
1983. Gamma-ray imaging with a rotating hexagonal uniformly redudant array. 158, 316

Czochralski, J.
1918. Ein neues Verfahren zur Messung der Kristallisationsgeschwindigkeit der Metalle. *Zeitschrift für Physikalische Chemie*, 92:219–221. 96

Daly, E., J. Lemaire, D. Heynderickx, and D. Rodgers
1996. Problems with models of the radiation belts. *IEEE Transactions on Nuclear Science*, 43(2):403–415. 63

De Leeuw, J. H.
1963. Electrostatic plasma probes. In *Fifth Biennial Gas Dynamics Symposium*, Pp. 1–44. 14, 79

Deptuch, G., G. Claus, C. Colledani, M. Deveaux, A. Gay, W. Dulinski, Y. Gornushkin, C. Hu-Guo, and M. Winter
2003. Development of monolithic active pixel sensors for charged particle tracking. *Nuclear Instruments and Methods in Physics Research A*, 511(1-2):240–249. 116

Dicke, R.
1968. Scatter-hole cameras for x-rays and gamma rays. *The Astrophysical Journal*, 153:L101 – L106. 15, 145, 146

Dioszegi, I., B. Yu, G. Smith, N. Schaknowski, J. Fried, P. E. Vanier, C. Salwen, and L. Forman
2013. A new pad-based neutron detector for stereo coded-aperture thermal neutron imaging. In *IEEE Nuclear Science Symposium Conference Record*. 147

Donnarumma, I., Y. Evangelista, R. Campana, J. in't Zand, M. Feroci, N. Lund, S. Brandt, J. Wilms, and C. Schmid
2012. The LOFT Wide Field Monitor simulator. *arXiv preprint*. 161

Dougherty, M.
2011. JUICE: exploring the emergence of habitable worlds around gas giants. Technical Report December. 67

Dowler, M., V. Aguero, and S. Sears
2002. Design and Test of a Solid State Charged Particle Detector for Cubesat. In *Proceedings of the 16th Annual/USU Conference on Small Satellites*. 130

Feister, I.
1950. Numerical evaluation of the Fermi beta-distribution function. *Physical Review*, 78(4):375–377. 275

Fenimore, E. E.
1978. Coded aperture imaging: predicted performance of uniformly redundant arrays. *Applied optics*, 17(22):3562–3570. 150

Fenimore, E. E. and T. M. Cannon
1978. Coded aperture imaging with uniformly redundant arrays. *Applied Optics*, 17(3):337–347. 150, 155, 174, 199

Fenimore, E. E. and T. M. Cannon
1981. Uniformly redundant arrays: digital reconstruction methods. *Applied optics*, 20(10):1858–1864. 156

Fermi, E.
1934. Tentativo di una Teoria Dei Raggi β. *Il Nuovo Cimento*, 11(1):1–19. 274

Finger, M. H. and T. A. Prince
1985. Hexagonal uniformly redundant arrays for coded-aperture imaging. *International Cosmic Ray Conference*, 3:295–298. 152, 153

Fossum, E.
1993. Active pixel sensors: are CCD's dinosaurs? *SPIE*, 1900:1–13. 15, 115

Funsten, H. O. and D. J. McComas
1998. Limited resource plasma analyzers: miniaturization concepts. In *Measurement techniques in space plasmas: particles*, Pp. 157–168. 91

Fürst, F., J. Wilms, R. E. Rothschild, K. Pottschmidt, D. M. Smith, and R. Lingenfelter
2009. Temporal variations of strength and location of the South Atlantic Anomaly as measured by RXTE. *Earth and Planetary Science Letters*, 281(3-4):125–133. 55

Gal, O., M. Gmar, O. P. Ivanov, F. Lainé, F. Lamadie, C. Le Goaller, C. Mahé, E. Manach, and V. E. Stepanov
2006. Development of a portable gamma camera with coded aperture. *Nuclear Instruments and Methods in Physics Research A*, 563(1):233–237. 147

Gal, O., C. Izac, F. Jean, F. Lainé, C. Lévêque, and A. Nguyen
2001. CARTOGAM–a portable gamma camera for remote localisation of radioactive sources in nuclear facilities. *Nuclear Instruments and Methods in Physics Research A*, 460(1):138–145. 147

Gal, O., C. Izac, F. Laine, and A. Nguyen
1997. CARTOGAM: a portable gamma camera. *Nuclear Instruments and Methods in Physics Research A*, 387(1):297–303. 147

Gatti, E. and P. Rehak
1984. Semiconductor drift chamber - An application of a novel charge transport scheme. *Nuclear Instruments and Methods in Physics Research*, 225(3):608–614. 111

Gatti, E., P. Rehak, and J. T. Walton
1984. Silicon drift chambers first results and optimum processing of signals. *Nuclear Instruments and Methods in Physics Research A*, 226(1):129–141. 15, 111

Gehrels, N., T. L. Cline, A. F. Huters, M. Leventhal, C. J. MacCallum, J. D. Reber, P. D. Stang, B. J. Teegarden, and J. Tueller
1985. A coded aperture imaging system optimized for hard X-ray and gamma-ray astronomy. *19th International Cosmic Ray Conference*, 3:305–306. 157

Geiger, H. and W. Müller
1928. Elektronenzählrohr zur Messung schwächster Aktivitäten. *Die Naturwissenschaften*, 16(31):617–618. 78

Ginet, G. P. and T. P. O'Brien
2009. AE-9 / AP-9 Trapped Radiation and Plasma Models Requirements Specification. Technical report. 21, 63, 64

Ginet, G. P., T. P. O'Brien, S. L. Huston, W. R. Johnston, T. B. Guild, R. Friedel, C. D. Lindstrom, C. J. Roth, P. Whelan, R. a. Quinn, D. Madden, S. Morley, and Y. J. Su
2013. AE9, AP9 and SPM: New models for specifying the trapped energetic particle and space plasma environment. 63

Glaser, D., J. Halekas, and P. Turin
2009. STEIN (SupraThermal Electrons, Ions and Neutrals), A New Particle Detection Instrument for Space Weather Research with CubeSats. In *23rd Annual AIAA/USU Conference on Small Satellites*, Pp. 1–9. 14, 91, 92, 93, 128

Gmar, M., M. Agelou, F. Carrel, and V. Schoepff
2011. GAMPIX: A new generation of gamma camera. *Nuclear Instruments and Methods in Physics Research A*, 652(1):638–640. 147, 313, 315

Goldstein, E.
1880. Ueber die Entladung der Electricität in verdünnten Gasen. *Annalen der Physik*, 247(13):833–856. 77

Goldwurm, A., P. David, L. Foschini, A. Gros, P. Laurent, A. Sauvageon, A. J. Bird, L. Lerusse, and N. Produit
2003. The INTEGRAL/IBIS scientific data analysis. *Astronomy and Astrophysics*, 411(1):L223–L229. 159, 160

Gombosi, T. I.
1998. *Physics of the Space Environment*. Cambridge University Press. 13, 14, 41, 45, 62

Gottesman, S. R. and E. E. Fenimore

1989. New family of binary arrays for coded aperture imaging. *Applied Optics*, 28(20):4344–4352. 152

Gunson, J. and B. Polychronopulos

1976. Optimum design of a coded X-ray telescope for rocket applications. *Monthly Notices of the Royal Astronomical Society*, 177:485–497. 146, 150

Gussenhoven, M. S., E. G. Müllen, and D. H. Bräutigam

1996. Improved understanding of the earth's radiation belts from the CRRES satellite. *IEEE Transactions on Nuclear Science*, 43(2 PART 1):353–368. 54

Hall, H.

1950. On the Evaluation of the Fermi β-Distribution Function. *Physical Review*, 79(4):745. 275

Hammond, C. M., W. C. Feldman, D. J. McComas, J. L. Phillips, and R. J. Forsyth

1996. Variation of electron-strahl width in the high-speed solar wind: ULYSSES observations. *Astronomy and Astrophysics*, 316:350–354. 48

Hargreaves, J. K.

1995. *The solar-terrestrial environment*, Cambridge Atmospheric and Space Science Series. Cambridge University Press. 52

Hell, N. and R. S. Bamberg

2010. *The Evolution of the South Atlantic Anomaly Measured by RHESSI*. Bachelor thesis. 56

Heynderickx, D.

2002. Radiation belt modelling in the framework of space weather effects and forecasting. *Journal of Atmospheric and Solar-Terrestrial Physics*, 64(16):1687–1700. 54

Hodgson, R.
1859. On a curious Appearance seen in the Sun. *Monthly Notices of the Royal Astronomical Society*, 20:15–16. 61

Hong, J., S. V. Vadawale, M. Zhang, E. C. Bellm, A. Yousef, J. Noss, J. E. Grindlay, and T. Narita
2004. Laboratory coded aperture imaging experiments: radial hole coded masks and depth-sensitive CZT detectors. *Proc. SPIE 5540, Hard X-Ray and Gamma-Ray Detector Physics VI*, 5540:10. 153, 316

Jedwab, J. and C. Mitchell
1988. Constructing new perfect binary arrays. *Electronics Letters*, 24(11):650–652. 152

Johnson, C. Y.
1969. Ion and neutral composition of the ionosphere. *Annals of the IQSY*, 5:197–213. 52

Johnstone, A. D.
1972. The geometric factor of a cylindrical plate electrostatic analyzer. *Review of Scientific Instruments*, 43(7):1030–1040. 80

Klein, C. A.
1968. Bandgap dependence and related features of radiation ionization energies in semiconductors. *Journal of Applied Physics*, 39(4):2029–2038. 14, 95

Klesh, A., J. Baker, J. Castillo-Rogez, L. Halatek, N. Murphy, C. Raymond, J. Bellardo, J. Cutler, and G. Lightsey
2013. Conference on Small Satellites. *27th Annual AIAA/USU Conference on Small Satellites*. 74

Knoll, G. F.
2000. *Radiation Detection and Measurement*, volume 3. Wiley. 102, 103, 106, 116

Kryuchkov, E. I. and A. K. Fedorenko
2012. Peculiarities of energy transport in the atmosphere by acoustic gravity waves. *Geomagnetism and Aeronomy*, 52(2):235–241. 53

Kurie, F. N. D., J. R. Richardson, and H. C. Paxton
1936. The radiations emitted from artificially produced radioactive substances. I. The upper limits and shapes of the β-ray spectra from several elements. *Physical Review*, 49(5):368–381. 275

Lee, A., R. Bedington, L. Comandar, A. Daurskikh, M. Hills, D. Hu, T. Nordheim, and R. Lee
2013. PoleCATS - a plasma instrumentation technology demonstration for REXUS-14. In *21st ESA Symposium on European Rocket and Balloon Programmes and Related Research*, Pp. 469–472. 296

Lei, F., P. R. Truscott, C. S. Dyer, B. Quaghebeur, D. Heynderickx, P. Nieminen, H. Evans, and E. Daly
2002. MULASSIS: A Geant4-based multilayered shielding simulation tool. *IEEE Transactions on Nuclear Science*, 49(6):2788–2793. 164

Lemaire, H., R. A. Khalil, K. Amgarou, J. C. Angélique, F. Bonnet, D. De Toro, F. Carrel, O. Giarmana, M. Gmar, N. Menaa, Y. Menesguen, S. Normand, A. Patoz, V. Schoepff, P. Talent, and T. Timi
2014. Implementation of an imaging spectrometer for localization and identification of radioactive sources. *Nuclear Instruments and Methods in Physics Research A*, 763:97–103. 147, 313, 315

Levine, A. M., H. Bradt, W. Cui, J. G. Jernigan, E. H. Morgan, R. Remillard, R. E. Shirey, and D. A. Smith
1996. First results from the all-sky monitor on the Rossi X-ray Timing Explorer. *The Astronomical Journal*, 469:33–36. 157

Lindstroem, G., M. Moll, and E. Fretwurst
1999. Radiation hardness of silicon detectors - a challenge from high-energy physics. *Nuclear Instruments and Methods in Physics Research A*, 426(1):1–15. 117

Llopart, X., M. Campbell, R. Dinapoli, D. San Segundo, and E. Pernigotti
2001. Medipix2, a 64-k pixel readout chip with 55 µm square elements working in single photon counting mode. In *Nuclear Science Symposium Conference Record*, volume 3, Pp. 1484–1488. IEEE. 129, 152

Longmire, C. and H. Brown
1949. Screening and relativistic effects on beta spectra. *Physical Review*, 75(2):264–270. 275

Lucy, L. B.
1974. An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79:745–749. 156

Luhmann, J. G.
1976. Auroral electron spectra in the atmosphere. *Journal of Atmospheric and Terrestrial Physics*, 38:605–610. 13, 57, 58

Lüke, H. D.
1987. Binäre und fast binäre orthogonale Folgen und Matrizen. *Frequenz*, 41(11-12). 152

MacWilliams, F. J. and N. J. Sloane
1976. Pseudo-Random Sequences and Arrays. *Proceedings of the IEEE*, 64(12):1715–1729. 152

Mannucci, A. J., B. D. Wilson, D. N. Yuan, C. H. Ho, U. J. Lindqwister, and T. F. Runge
1998. A global mapping technique for GPS-derived ionospheric total electron content measurements. *Radio Science*, 33(3):565. 70

Marleau, P., J. Brennan, E. Brubaker, and J. Steele
2010. Results from the coded aperture neutron imaging system. *IEEE Nuclear Science Symposium & Medical Imaging Conference*, Pp. 1640–1646. 147

McComas, D. J., S. J. Bame, and P. Barker
1998. Solar wind electron proton alpha monitor (SWEPAM) for the Advanced Composition Explorer. *Space Science Reviews*, 86:563–612. 13, 44

McFadden, J. P., C. W. Carlson, M. H. Boehm, and T. J. Hallinan
1987. Field-aligned electron flux oscillations that produce flickering aurora. *Journal of Geophysical Research*, 92(A10):11133. 20, 301, 302

McFarland, R. L.
1973. A family of difference sets in non-cyclic groups. *Journal of Combinatorial Theory, Series A*, 15(1):1–10. 152

Meikle, S. R., R. R. Fulton, S. Eberl, M. Dahlbom, K. P. Wong, and M. J. Fulham
2001. An investigation of coded aperture imaging for small animal SPECT. *IEEE Transactions on Nuclear Science*, 48(3 II):816–821. 147

Mende, S. B., S. E. Harris, H. U. Frey, V. Angelopoulos, C. T. Russell, E. Donovan, B. Jackel, M. Greffen, and L. M. Peticolas
2008. The THEMIS array of ground-based observatories for the study of auroral substorms. *Space Science Reviews*, 141(1-4):357–387. 70

Mertz, L. and N. O. Young
1961. Fresnel Transformation of Images. In *Optical Instruments and Techniques (Chapman and Hall Ltd)*, P. 305. 145, 148

Mitchell, E. F., H. M. Araújo, E. Daly, N. Guerrini, S. Gunes-Lasnet, D. Griffin, A. Marshall, A. Menicucci, T. Morse, O. Poyntz-Wright, R. Turchetta, and S. Woodward
2014. The Highly Miniaturised Radiation Monitor. *Journal of Instrumentation*, 9:17. 14, 88, 90, 93

Mohammadzadeh, A., H. Evans, P. Nieminen, E. Daly, P. Vuilleumier, P. Bühler, C. Eggel, W. Hajdas, N. Schlumpf, A. Zehnder, J. Schneider, and R. Fear
2003. The ESA Standard Radiation Environment Monitor Program First Results from PROBA-I and INTEGRAL. *IEEE Transactions on Nuclear Science*, 50(6):2272–2277. 93

Mott-Smith, H. M. and I. Langmuir
1926. The theory of collectors in gaseous discharges. *Physical Review*, 28(4):727–763. 79

Munakata, R.
2009. Cubesat design specification rev. 12. *The CubeSat Program, California Polytechnic State University*. 72

Noble, P.
1968. Self-scanned silicon image detector arrays. *IEEE Transactions on Electron Devices*, 15(4):202–209. 115

Ogasawara, K., K. Asamura, T. Takashima, Y. Saito, and T. Mukai
2006. Rocket observation of energetic electrons in the low-altitude auroral ionosphere during the DELTA campaign. *Earth Planets Space*, 58:1155–1164. 57

Owens, A. and A. Peacock
2004. Compound semiconductor radiation detectors. In *Nuclear Instruments and Methods in Physics Research A*, volume 531, Pp. 18–37. 95

Parker, E. N.
1958. Dynamics of the Interplanetary Gas and Magnetic Fields. *The Astrophysical Journal*, 128:664. 43

Paschmann, G., H. Loidl, P. Obermayer, M. Ertl, R. Laborenz, N. Sckipke, W. Baumjohann, C. W. Carlson, and D. W. Curtis
1985. The Plasma Instrument for AMPTE IRM. *Geoscience and Remote Sensing, IEEE Transactions on*, GE-23(3):262–266. 84

Pierrard, V., M. Maksimovic, and J. Lemaire
2001. Selfconsistent model of solar wind electrons. *Journal of Geophysical Research*, 106(A12):29305–29312. 47, 304

Pinsky, L., S. M. Hoang, J. Idarraga-Munoz, M. Kroupa, N. Stoffle, A. Bahadori, E. Semones, J. Jakubek, Z. Vykydal, D. Turecek, S. Pospisil, H. Kitamura, and S. Kodaira
2014. Summary of the First Year of Medipix-Based Space Radiation Monitors on the ISS. *Aerospace Conference, 2014 IEEE*, Pp. 1–8. 129

Pinsky, L. S., A. Empl, S. Hoang, N. Stoffle, J. Jakubek, Z. Vykydal, D. Turecek, S. Pospisil, H. Kitamura, O. Ploc, Y. Uchihori, Y. Nakahiro, C. Amberboy, J. Hauss, K. Lee, E. Semones, N. Zapp, R. Parker, and D. Cooke
2012. Preparing for the first Medipix detectors in space. In *IEEE Aerospace Conference Proceedings*, Pp. 1–6. 129

Pizzo, V. J.
1978. A three-dimensional model of corotating streams in the solar wind I - Theoretical foundations. *NASA Technical Memorandum 79580*. 13, 46

Potuluri, P., U. Gopinathan, J. Adleman, and D. Brady
2003. Lensless sensor system using a reference structure. *Optics Express*, 11(8):965–974. 153

Pratt, R., H. Tseng, C. Lee, L. Kissel, C. MacCallum, and M. Riley
1977. Bremsstrahlung energy spectra from electrons of kinetic energy 1 keV $\leq$ T1 $\leq$ 2000 keV incident on neutral atoms 2 $\leq$ Z $\leq$ 92. *Atomic Data and Nuclear Data Tables*, 20(2):175–209. 106

Priest, E. and T. Forbes
2000. Magnetic reconnection : MHD theory and applications. *Cambridge University Press*. 43

Richardson, W.
1972. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1):55–59. 156

Rishbeth, H. and P. J. S. Williams
1985. The EISCAT ionospheric radar - The system and its early results. *Royal Astronomical Society, Quarterly Journal*, 26:478–512. 69

Roques, J. P.
1987. Fast decoding algorithm for uniformly redundant arrays. *Applied optics*, 26(18):3862–5. 156

Russell, C. T.
1993. Planetary magnetospheres. *Reports on Progress in Physics*, 56(6):687–732. 66

Russell, C. T., P. J. Chi, D. J. Dearborn, Y. S. Ge, B. Kuo-Tiong, J. D. Means, D. R. Pierce, K. M. Rowe, and R. C. Snare
2008. THEMIS ground-based magnetometers. *Space Science Reviews*, 141(1-4):389–412. 69

Rutherford, E.
1899. Uranium radiation and the electrical conduction produced by it. *Philosophical Magazine*, 47(284):109–163. 78

Sambo, L., J. B. Stephen, S. Bonettini, G. Zanghirati, and F. Frontera
2009. Improving the angular resolution of coded aperture instruments using a modified Lucy-Richardson algorithm for deconvolution. In *Proceedings of Science*. 156, 200

Sawyer, D. M. and J. I. Vette
1976. AP-8 trapped proton environment for solar maximum and solar minimum. *NASA STI/Recon Technical Report N*, 77. 63

Schiller, Q. G., A. Mahendrakumar, and X. Li
2010. REPTile : A Miniaturized Detector for a CubeSat Mission to Measure Relativistic Particles in Near-Earth Space. In *USU Conference on Small Satellites*, Pp. SSC10–VIII–1. 14, 88, 89, 93

Schwabe, H.
1844. Sonnenbeobachtungen im Jahre 1843. *Astronomische Nachrichten*, 21:233. 61

Schwenn, R.
2005. Space weather: The solar perspective. *Living Reviews in Solar Physics*, 3(1). 48

Seltzer, S. M. and M. J. Berger
1985. Bremsstrahlung spectra from electron interactions with screened atomic nuclei and orbital electrons. *Nuclear Instruments and Methods in Physics Research B*, 12(1):95–134. 106

Shockley, W.
1949. The theory of p-n junctions in semiconductor and p-n junction transistors. *Bell System Technical Journal*, Pp. 435–489. 99

Shockley, W.

1950. *Electrons and holes in semiconductors*, 7 edition. Princeton: D. Van Nostrand. 96, 97

Siegl, M.

2009. *Standard radiation environment monitor: simulation and inner belt flux anisotropy investigation*. Master thesis. 14, 87, 93

Siklér, F.

2012. A parametrization of the energy loss distributions of charged particles and its applications for silicon detectors. *Nuclear Instruments and Methods in Physics Research A*, 691:16–29. 107

Siklér, F. and S. Szeles

2012. Optimized estimation of energy loss rate for charged particles from energy deposit measurements in tracking detectors. *Nuclear Instruments and Methods in Physics Research A*, 687:30–39. 107

Singer, J.

1938. A theorem in finite projective geometry and some applications to number theory. *Transactions of the American Mathematical Society*, 43(3):377–377. 150

Skinner, G. and J. Grindlay

1993. Coded masks with two spatial scales. *Astronomy and Astrophysics*, 276:673–681. 158

Skinner, G. K.

2004. Coded mask imagers when to use them - And when not. *New Astronomy Reviews*, 48(1-4):205–208. 132, 140

Skinner, G. K.

2008. Sensitivity of coded mask telescopes. *Applied Optics*, 47(15):2739–2749. 149

Skorokhod, T. V. and G. V. Lizunov
2012. Localized packets of acoustic gravity waves in the ionosphere. *Geomagnetism and Aeronomy*, 52(1):88–93. 52, 59

Smith, A., T. Beuselinck, J. Dalsgaard Nielsen, J. De Keyser, A. Gregorio, D. Kataria, V. Lappas, F. Lubken, J. Moen, S. Palo, R. Reinhard, A. Ridley, J. Rotteveel, G. Schmidtke, and T. Schmiel
2012. Sensor selection working group - Final Report. Technical report, QB50 Consortium. 73

Sosin, Z.
2012. Description of the plasma delay effect in silicon detectors. *Nuclear Instruments and Methods in Physics Research A*, 693:170–178. 173

Spieler, H.
2005. *Semiconductor Detector Systems*. Oxford University Press. 14, 15, 98, 99, 100, 101, 116, 120

Spieler, H.
2012. Electronics and data acquisition. *Nuclear Instruments and Methods in Physics Research A*, 666:197–222. 15, 113, 114

Stearns, D. G. and J. D. Wiedwald
1989. Response of charge-coupled devices to direct electron bombardment. *Review of Scientific Instruments*, 60(6):1095–1103. 265

Stern, R. A., L. Shing, and M. M. Blouke
1994. Quantum efficiency measurements and modeling of ion-implanted, laser-annealed charge-coupled devices: x-ray, extreme-ultraviolet, ultraviolet, and optical data. *Applied Optics*, 33(13):2521–2533. 265

Stern, R. A., L. Shing, P. R. Catura, M. D. Morrison, D. W. Duncan, J. R. Lemen, T. Eaton, P. J. Pool, R. Steward, D. M. Walton, and A. Smith
2004. Characterization of the Flight CCD Detectors for the GOES N and O Solar X-ray Imagers. *Proceedings of SPIE Vol. 5171 (Telescopes and Instrumentation for Solar Astrophysics)*, Pp. 77–88. 225, 283

Sternheimer, R. M.
1956. Density effect for the ionization loss in various materials. *Physical Review*, 103(3):511–515. 103

Sternheimer, R. M., S. M. Seltzer, and M. J. Berger
1982. Density effect for the ionization loss of charged particles in various substances. *Physical Review B*, 26(11):6067–6076. 103

Talebitaher, A., P. M. Shutler, S. V. Springham, R. S. Rawat, and P. Lee
2012. Coded aperture imaging of alpha source spatial distribution. *Radiation Measurements*, 47(10):992–999. 146

Taylor, B., C. Underwood, A. Dyer, C. Ashton, S. Rason, and J. Browning
2011. The micro radiation environment monitor (MuREM) and SSTL radiation monitor (SSTL RM) on TechDemoSat-1. In *Radiation and Its Effects on Components and Systems (RADECS)*, Pp. 535–540. 87, 93

Taylor, B., C. Underwood, A. Dyer, C. Ashton, S. Rason, and J. Browning
2012. The Micro Radiation Environment Monitor (MuREM) and SSTL Radiation Monitor (SSTL RM) on TechDemoSat-1. *IEEE Transactions on Nuclear Science*, 59(4):1060–1065. 130

Teal, G. K. and J. B. Little
1950. Growth of Germanium Single Crystals. *Physical Review*, 78(5):647. 96

Teal, G. K., M. Sparks, and E. Buehler
1951. Growth of germanium single crystals containing p-n junctions. *Physical Review*, 81(4):637. 97

Teegarden, B., T. Cline, N. Gehrels, G. Porreca, and J. Tueller
1985. The gamma-ray imaging spectrometer (GRIS): a new balloon-bourne experiment for gamma-ray astronomy. *19th Intern. Cosmic Ray Conf.*, 3:307–310. 157

Thomson, J. J.
1897. XL. Cathode Rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):293–316. 77

Torsti, J., E. Valtonen, M. Lumme, P. Peltonen, T. Eronen, M. Louhola, E. Riihonen, G. Schultz, M. Teittinen, K. Ahola, C. Holmlund, V. Kelhä, K. Leppälä, P. Ruuska, and E. Strömmer
1995. Energetic particle experiment ERNE. *Solar Physics*, 162(1-2):505–531. 82

Vampola, A. L.
1998. Measuring energetic electrons - what works and what doesn't. In *Measurement techniques in space plasmas: particles*, Pp. 339–355. 109

Vampola, A. L.
2000. The hazardous space particle environment. *IEEE Transactions on Plasma Science*, 28(6):1831–1839. 61

Van Allen, J. A.
1958. Observation of High Intensity Radiation by Satellites 1958 Alpha and Gamma. *Journal of Jet Propulsion*, 28(9):588–592. 53, 78

Van Allen, J. A.
1959. The geomagnetically trapped corpuscular radiation. *Journal of Geophysical Research*, 64(11):1683–1689. 54, 78

Vanier, P. E.
2003. Improvements in coded aperture thermal neutron imaging. In *Optical Science and Technology, SPIE's 48th Annual Meeting. International Society for Optics and Photonics*. 147

Venkatarangan, P. and D. McEwen
1979. Electron measurements (18 keV - 20 eV) in auroral events. *Planetary and Space Science*, 27(5):669–677. 57

Vette, J. I.
1991. The AE-8 trapped electron model environment. Technical report. 63

Walton, D., P. Thomas, J. Culhane, B. Jordan, A. Smith, A. Dibbens, and L. Bradley
2003. The CCD and readout electronics for the OMC instrument on Integral. *Astronomy and Astrophysics*, 411:275–279. 226

Whyntie, T. and M. Harrison
2015. Full simulation of the LUCID experiment in the Low Earth Orbit radiation environment. *Journal of Instrumentation*, 10(03):C03043–C03043. 129, 131

Widenhorn, R., M. M. Blouke, A. Weber, A. Rest, and E. Bodegom
2002. Temperature dependence of dark current in a CCD. *Proc. SPIE Vol. 4669*, 4669(April):193–201. 109

Wilson, F. L.
1968. Fermi's Theory of Beta Decay. *American Journal of Physics*, 36(12):1150–1160. 274

WolframAlpha
2016. WolframAlpha. *www.wolframalpha.com*. 51

Wright, D.
2011. Geant4 Physics Reference Manual. 166